The Reasonable and the Relevant: Legal Standards of Proof

Georgi Gardiner
University of Tennessee

This essay appears in *Philosophy & Public Affairs*

Abstract. According to a common conception of legal proof, satisfying a legal burden requires establishing a claim to a numerical threshold. Beyond reasonable doubt, for example, is often glossed as 90% or 95% likelihood given the evidence. Preponderance of evidence is interpreted as meaning at least 50% likelihood given the evidence. In light of problems with the common conception, I propose a new 'relevant alternatives' framework for legal standards of proof. Relevant alternative accounts of knowledge state that a person knows a proposition when their evidence rules out all relevant error possibilities. I adapt this framework to model three legal standards of proof—the preponderance of evidence, clear and convincing evidence, and beyond reasonable doubt standards. I describe virtues of this framework. I argue that, by eschewing numerical thresholds, the relevant alternatives framework avoids problems inherent to rival models. I conclude by articulating aspects of legal normativity and practice illuminated by the relevant alternatives framework.

Keywords. Legal standards of proof, relevant alternatives theory, proof paradox, beyond reasonable doubt, preponderance of evidence

1. The Limits of Balance

A common conception holds that epistemic justification is a matter of probabilistic likelihood given the evidence. On this widely-held view, a claim is epistemically justified if it meets a threshold of probabilistic likelihood given the evidence; if, given a body of evidence, claim p is more likely than claim q, this means that p is more epistemically justified than q by that evidence.[1] Against this background, theorists discuss the probabilistic threshold for the legal standard of proof 'beyond reasonable doubt'.

Beyond reasonable doubt is the epistemic threshold required for criminal conviction in increasingly many legal systems, including most anglophone countries, Germany, Sweden, Italy, and Israel.[2] Theorists argue that—given the seriousness of criminal proceedings, the distinctive harms of false conviction, and the importance of public trust in the system—the relevant standard of proof must be a high threshold. Accordingly beyond reasonable doubt is often glossed as around 90% or 95% probability of guilt given the evidence. Debates rage about whether this figure is too low or high.[3] Theorists also debate whether the likelihood should be sensitive to practical stakes involved. Laudan argues, for instance, that high costs of false acquittal for violent crime means that the standard of proof governing conviction for violent crime should be lowered and suggests a threshold in the vicinity of 56% to 67%.[4] Others suggest that, given the high costs of false conviction, a standard of 90% or 95% likelihood is not sufficiently demanding.

---

1. See Russell (1948: ch. 6), Chisholm (1957: 28), Goldman (1986: §5.5), Fumerton (1995), Pryor (2004: 350f), and BonJour (2010). See also Smith (2010; 2016). For detailed overviews about whether legal standards of proof amount to probabilistic thresholds, see Di Bello (2013), Ho (2015), and Gardiner (2019).
2. Mulrine (1997), *Winship*, 397 U.S. at 371–72. Although Roberts (ms) questions the prevalence of the beyond reasonable doubt standard.
3. There are two related debates. Some focus on whether 'beyond reasonable doubt', typically understood as around 90% or 95% likelihood, ought to govern criminal conviction. More commonly, theorists focus on whether numerical figures around 90% or 95% are the correct range to capture 'beyond reasonable doubt'.
4. Laudan (2011: 207). For rebuttals, see Risinger (2010) and Gardiner (2017a).

The picture painted by this commonly-held conception is that levels of epistemic justification can be understood as a scale of quantified likelihood given the evidence. Probability given the evidence is known as the 'balance' of the evidence; it depends on the balance between unfavourable and favourable evidence. We can call this conception of epistemic justification the 'quantifiable balance' conception: Epistemic justification is simply a matter of quantifiable evidential balance.

This picture has virtues. It is simple. By allowing quantificational representation, it renders epistemic justification amenable to mathematical treatments, such as Bayesianism. And it gets almost every case right: If p is more likely than q, given the evidence, then in almost every real-life case, p is indeed more epistemically justified than q.[5]

But this picture has weaknesses. To see why, consider the following case. It is 9pm Eastern Time on the night of the US midterm elections. Some voting places have closed, and some results have been announced. Sarah cannot watch the results come in, but she checks a respected political analysis website. The site estimates the Democrats have a 75% chance of winning a House majority. Sarah supports the Dems, and so is pleased. She reads that their estimate is largely based on survey data, bolstered by extensive insight into social trends, and the political climate. They used the best political surveys available when voting places opened that morning.

Sarah checks a second respected website. Like the first, this website studies the social and political climate and the recent survey data. It also has access, she notes, to mail ballot results, voter turnout throughout the day, various real-time vote tallies, which seats are already called, and exit surveys, which are more reliable than ordinary political surveys. The second website estimates the Dems have a 65% chance of winning the House. Sarah is even more pleased.

The second estimate says her desired result is less probable, given the evidence. So why is Sarah more pleased? In this paper I explain why Sarah's response can be rational. The estimated likelihood given the evidence—the evidential balance—is lower. If this were all that mattered Sarah should be less happy. But something else has clearly increased: The amount of evidence available. The second estimate is on a substantially better footing.[6]

As John Maynard Keynes (1921: 71) wrote,

> As the relevant evidence at our disposal increases, the magnitude of the probability of the argument may either decrease or increase, according as the new knowledge strengthens the unfavourable or the favourable evidence; but something seems to have increased in either case,—we have a more substantial basis upon which to rest our conclusion. I express this by saying that an accession of new evidence increases the weight of an argument. New evidence will sometimes decrease the probability of an argument, but it will always increase its 'weight'.

In this essay I draw on the relevant alternatives theory of knowledge to illuminate Keynes's insight and elucidate the change in Sarah's epistemic position. In short, new evidence can make a claim more likely—evidence can affect epistemic balance—but this does not exhaust the epistemic force of new evidence. Evidence can also address relevant error possibilities. I argue that relevant alternatives theory yields a framework to model legal standards of proof in a non-numerical way.[7]

I first sketch a relevant alternatives theory of knowledge and explain how it differs from the quantifiable balance conception of evidential support. In section three I employ the relevant

[5] Buchak (2014) and Smith (2010) propose counterexamples.

[6] In general, two evidence sets might both support a likelihood of n%. But if one contains significantly more evidence, then a resulting judgement can be more justified. The distinction between evidential balance and amount of evidence is not captured by the simple 'quantifiable balance' model of justification. Once the conception is amended to capture this distinction, the above-cited virtues of simplicity and amenability to mathematical treatments are threatened.

[7] This essay employs the relevant alternatives framework to model legal standards of proof. A related project harnesses the framework to model evidential weight: Namely, evidential weight is a measure of which error possibilities are ruled out by the evidence.

alternatives framework to model three legal standards of proof: The preponderance of evidence, clear and convincing evidence, and beyond reasonable doubt standards. Finally, I describe virtues of the proposal.

2. The Relevant Alternatives Account of Knowledge

Relevant alternatives theory was first developed as an account of knowledge.[8] Suppose Matt, who knows much about animals, looks into a pen at the zoo. He sees a black and white striped equid. Teaching animals to his daughter, Matt accurately reports 'there is a zebra in the pen'. Intuitively Matt knows the animal is a zebra. Matt can distinguish zebras from other animals, such as camels, yaks, and deer. He would not, furthermore, mistake an okapi or zebroid for a zebra.

But Matt cannot rule out every non-zebra possibility. Consider the possibility that the animal in the pen is a cleverly disguised mule. Since Matt lacks the capacity to distinguish zebras from cleverly disguised mules, his evidence cannot rule out this error possibility. But, the thought goes, this possibility is outlandish and can be disregarded. In ordinary cases this remote possibility is not relevant to whether Matt knows the animal is a zebra.[9]

Now consider Tim, who knows less about animals. He takes his daughter around the zoo. Keen to teach her, he points into the pen and says 'there is a zebra in the pen'. Tim can discriminate zebras from most other animals, such as camels, yaks, and deer. But Tim is ignorant of okapis and zebroids, and cannot visually discriminate these from zebras. Since zoos typically include such animals, the okapi and zebroid possibilities are relevant. Given the relevant unaddressed error possibilities, Tim does not know the animal is a zebra. His evidence cannot rule out relevant alternatives.

Suppose Tim looks at the sign, which says 'zebra'. Fortified by this additional evidence, Tim's total evidence does address the relevant alternative that the animal is an okapi, zebroid, or other non-zebra. Given Tim sees the sign, he knows. There are still some error possibilities that Tim's total evidence does not eliminate, but they are relatively remote. An example is that the animal in the pen is one he cannot visually distinguish from a zebra when he sees it, and—given he observes the sign—the sign erroneously says 'zebra'. This is an error possibility—a subset of the non-zebra possibilities—that Tim cannot rule out.[10] But it is relatively marginal, and plausibly can be ignored. Ambivalence about whether this error possibility can be disregarded, furthermore, plausibly co-varies with ambivalence about whether Tim knows.

Reflecting on Matt and Tim suggests the following condition on knowledge.

> Relevant alternatives condition for knowledge: S knows that p only if S can rule out relevant alternatives to p.

This condition has three main components: Alternatives, relevance, and ruling out. I sketch each in turn. An alternative, also known as an 'error possibility', is a proposition incompatible with the target belief. If the alternative obtains, then the belief is false; if the object is a disguised mule, it is not a zebra. There are many alternatives to p. These include that the object is a disguised mule, lifelike robot, okapi, zedonk, pig, or zoo keeper, and that the fathers hallucinate near an empty pen. Each of these possibilities admit of myriad sub-alternatives. These include, for example, the sub-alternatives incorrectly-labelled okapi and correctly-labelled okapi. These alternatives themselves can be

---

[8] Early influential treatments include Austin (1946), Dretske (1970), Stine (1976), Goldman (1976), Cohen (1986), and Lewis (1996). More recently, see Ho (2008), Lawlor (2013), McKinnon (2013), Amaya (2015), Gerken (2017), Ichikawa (2017), and Moss (2018a, 2018b). Rysiew (2006) and Bradley (2014) emphasise the plausibility and universal appeal of a relevant alternatives condition, and articulate how the condition is consistent with a wide range of epistemological views.

[9] Note, though, such shenanigans occur: www.bbc.co.uk/news/world-africa-44968509

[10] As I describe below, an error possibility is best understood as a collection of sub-alternative possibilities.

subdivided into sub-alternatives. There are many ways the sign could be mislabelled, for example, including accident, deceit, and pen-cleaning expedient. In general, alternatives have a branching structure, constituted by increasingly specific possibilities. Some of these alternatives are inconsistent with the fathers' evidence and some can be dismissed as irrelevant. This leads us to the second component, relevance.

Relevant alternatives are those error possibilities that cannot be properly ignored. We have some intuitive grasp on which alternatives are relevant for knowledge attributions. The possibility that the animal is a disguised mule can, in ordinary circumstances, be properly ignored. It is outlandish. The possibility that the animal is an okapi cannot be properly ignored. To know the animal is a zebra, a person must be able to address this possibility. According to the relevant alternatives condition on knowledge, we have an intuitive sense of which error possibilities are relevant. If the person's evidence eliminates all relevant alternatives, then we ascribe knowledge to that person. If there are relevant alternatives the person's evidence cannot address, then we deny the person possesses knowledge.

Perhaps in some cases we are uncertain whether the error possibility can be properly ignored. This does not threaten the relevant alternatives account if such uncertainty tends to correspond to uncertainty about whether knowledge attributions are appropriate.

David Lewis (1996) provides some guidance about when error possibilities cannot be properly ignored. His account, which I partially sketch, focuses on error possibilities relevant to knowledge ascriptions. He claims, for example, that alternatives are relevant if they are true. If the animal is an okapi, then this error possibility is relevant. Secondly, an error possibility cannot be properly ignored if it is believed or is being considered by the subject. If Matt wonders whether the animal is an okapi, for example, then this error possibility is relevant. Finally, an error possibility cannot be properly ignored if it saliently resembles an error possibility that cannot be properly ignored. If the possibility that the animal is a zedonk is relevant, perhaps because Linda considers this possibility, and zedonks saliently resemble zorses, then the possibility that the animal is a zorse is also relevant.[11]

As we shall see, Lewis's specific account of relevance will not suit our purposes. On Lewis's account, which alternatives are relevant is fickle. It can depend on what people are currently considering. If someone considers a skeptical possibility, such as a clever illusion, it becomes relevant. Since a zebra illusion cannot be ruled out by the fathers' evidence, knowledge cannot be attributed. Legal standards of proof are not hostage to how creative people are at devising error possibilities consistent with the evidence. And legal proof does not disappear and reappear depending on conversational context. But even if we depart from Lewis's account of relevance, the framework is a helpful starting point. It provides an intuitive picture: Some error possibilities are farfetched, they can be properly ignored. Others cannot be properly ignored—they seem important, relevant, reasonable—they must be ruled out.

This brings us to the third condition, ruling out. The nature of ruling out is also controversial. One way to gloss ruling out an alternative is possessing some evidence that discriminates the truth from the alternative. To know the animal is a zebra, Matt must possess evidence differentiating the truth from relevant alternatives, such as the animal's being an okapi. Perhaps he can tell the animal's stripes continue across the torso, for instance, which is a feature okapis lack. Without such evidence, Matt cannot eliminate the relevant error possibility and so lacks knowledge. Evidence that rules out its being an okapi need not be wholly infallible. It is possible the animal is an okapi with abnormal colouring. This sub-alternative of the okapi error possibility is compatible with Matt's evidence. But such remote error sub-possibilities are properly ignored.

Given that alternatives have a branching structure, error possibilities are typically not wholly eliminated outright. Instead they are split by new evidence. Some sub-possibilities are eliminated, and

---

[11] Lewis (1996: 556) outlines constraints on the resemblance condition.

others are farfetched and can be ignored. For a remaining sub-possibility to undermine a knowledge attribution, it must be both within the sphere of relevance and not eliminated by the evidence.

Since error possibilities are typically divisible, it can be helpful to talk of error possibilities being 'addressed' by the evidence, rather than eliminated by the evidence. Alternatives are not strictly ruled out, since some sub-alternatives are compatible with the evidence. There will always be creative skeptical possibilities. But remaining error sub-possibilities are increasingly ignorable as they are divided into increasingly fanciful and marginal possibilities. Thus we can say an error possibility is 'addressed' by the evidence when the possibility is rendered into branched alternatives, and each node is either incompatible with the evidence or lies beyond the threshold of what qualifies as relevant. Strictly speaking, then, error possibilities are addressed rather than eliminated, since in almost every case some remaining remote error sub-possibilities are consistent with the evidence. (This inevitable remainder fuels the skeptical challenge.[12])

Note that addressing error possibilities is normal and automatic. It is something we do all the time, without thinking about it. Our evidence always rules out uncountably many alternatives, such as there being a tiger in the room or a spoon on the desk. If challenged in everyday life, we appeal to our evidence: 'Of course it's a zebra, look at the stripes.' We need not think about error possibilities explicitly, or think of ourselves as eliminating possibilities.

Evidence that rules out more error possibilities typically also makes a claim more statistically likely; typically a claim becomes increasingly likely given the evidence as the only remaining uneliminated error possibilities are increasingly obscure, remote, outlandish, or convoluted. Eventually remaining uneliminated sub-alternatives are like this: The animal is an okapi… in a mislabelled pen, standing at a funny angle so its neck appears truncated, with abnormal colouring… and a mohican mane wig… This possibility is both statistically unlikely and remote. But sometimes eliminating error possibilities and probabilification come apart.

Lotteries provide a classic example of evidence that makes a claim statistically likely yet leaves uneliminated a relevant error possibility. Reflecting on the statistical evidence, Lottie believes her ticket will not win the 500-ticket lottery. Her belief is very likely true, given the evidence. But her evidence is unable to address a relevant alternative, namely that her ticket wins. As Lewis (1996) argues, a relevant alternatives condition on knowledge explains the intuition that Lottie fails to know her ticket has not won.

For evidence that addresses many error possibilities but fails to significantly affect evidential balance, consider masterful crime fiction before the 'big reveal' denouement. The reader receives many suggestive clues, but nonetheless the perpetrator is far from obvious. There are three main suspects. If the butler did it, the murder must have occurred before 3pm. So the corpse must have been in the drawing room when the son entered to collect his jacket. How did he fail to notice? It might have been the dairymaid. But only if the farm boy is lying or confused about where she was that afternoon. And only if the deceased accepted drink from her. But why would he do that? Or it could have been the sweep. But only if his bag contained clean clothes and only if he passed through the kitchen twice without being noticed.

The many pieces of evidence chop the error possibilities into tantalising sub-alternatives. Some are eliminated by the evidence, such as the sub-alternative where the dairymaid is the killer yet the farm boy's report is accurate.[13] Others are too fanciful and need not be taken seriously, such as that

[12] It is no coincidence that the *Cogito* is an exception. There are no uneliminated error sub-possibilities; the skeptical challenge cannot root there.

[13] Or, more carefully, this sub-alternative is *addressed* by the evidence, since there might be further sub-possibilities consistent with the boy telling the truth. Perhaps the dairymaid used a chemical to thwart forensic testing of the time

she has a secret twin who has been hidden from the villagers since birth. Other possibilities remain—not ruled out, yet not outlandish—such as that she is the killer and the farm boy is covering for her. An ideal mystery novel might have a few competing hypotheses that are approximately balanced. This exemplifies evidence having insignificant balance—the probabilities are close to even—yet significant weight; many sub-alternatives and sub-sub-alternatives are addressed.

Plausibly this feature explains the appeal of Season One of the hit podcast *Serial*. The overall probability of guilt and innocence seems relatively well-balanced, but the myriad pieces of evidence sub-divide the error possibility that Adnan Syed is innocent into increasingly narrow sub-alternatives. Perhaps some of these uneliminated sub-alternatives, although finely sliced, can be reasonably entertained. Perhaps, in other words, there remains some reasonable doubt.

3. Three Legal Standards

The relevant alternatives framework offers the following schema. There is a claim, p, such as that the pen contains a zebra. There are various not p possibilities, which can typically be sub-divided. Some of these sub-alternatives are eliminated by the evidence. Others lie outwith a particular threshold; that is, they are remote enough to properly ignore.

Lewis posits the 'relevant to knowledge' threshold. If an error possibility is relevant to knowledge—and is consistent with the person's evidence—the person fails to satisfy the condition for knowledge. Lewis claims that whether the error possibility can be properly ignored depends on features of conversational context and can accordingly change quickly.

We can set aside the conditions for knowledge, and instead employ the relevant alternatives framework to illuminate three legal standards of proof: Preponderance of evidence, clear and convincing evidence, and beyond reasonable doubt standards. In short, these three standards require addressing increasingly farfetched error possibilities. They mark three concentric circles, which are thresholds of remoteness for error possibilities. The remoteness of an error possibility and the location of the three thresholds are stable. Which error possibilities can be properly ignored cannot fluctuate with conversational context. (I return to what determines the thresholds below.) In what follows I first describe the scaffolding of the view. I then develop an example to illustrate.

Consider a claim such as:

> R: The defendant made an agreement with the complainant.

This claim is part of a broader set of claims—some contested, some not—in a legal case. If the claims are true, the defendant is legally liable. The complainant alleges that R; the defendant denies. There are many error possibilities—many ways the claim might be false. Some sub-possibilities are ordinary and mundane, such as the complainant misunderstood their communications or was too presumptuous and hasty. Some are farfetched and remote: A third party imitated the defendant daily for years in order to falsely instil belief R in the complainant.

On this proposal, the possibilities can be classed into three concentric circles; three tiers of disregardability. The closest sphere encloses all preponderant error possibilities. Preponderant error possibilities are the most important and significant, and so the least disregardable. 'Preponderance of the evidence' is often interpreted as meaning at least 50% statistical likelihood given the evidence. This

---

of death. But such sub-sub-possibilities are outlandish and can be ignored. (And if not properly ignored, they must be ruled out.)

corresponds to a quantitative reading of 'preponderance', suggested by a numerical gloss on 'most'. But the terms 'most' and 'preponderance' yield both a quantitative and a qualitative reading.[14]

Consider expressions like 'most beautiful beaches win awards' and 'most dear friends attended the funeral'. These expressions can be read two ways. On the quantitative reading there are a number of beautiful beaches and dear friends. More than half of them received an award or attended the funeral. The qualitative reading, which can be prompted by adding 'the' to the beginning of the expressions, says nothing about any proportions or majorities. Instead it suggests that some of the beaches are particularly gorgeous and some of the friends are intimately close. They are the ones that win awards and attended the funeral. I urge a similar interpretation for 'preponderance'. A quantitative reading forces the probabilistic interpretation of 'at least 50% likely given the evidence'. But the qualitative reading creates room for a non-numerical gloss on the legal standard.

The second tier marks those error possibilities that must be addressed in order to satisfy the clear and convincing evidence standard. Sub-alternatives in this sphere are more remote. Thus evidence might satisfy the preponderance standard, by addressing the most preponderant error possibilities, but fail to satisfy the more demanding standard.

Perhaps the error possibilities within this sphere match those relevant to knowledge. That is, perhaps possessing knowledge and satisfying the 'clear and convincing evidence' standard require addressing the same error possibilities. This proposal assumes, of course, that knowledge is not sensitive to fugacious features of conversational context, such as mentioning fanciful or skeptical error possibilities. Some motivation for endorsing this relation to knowledge is that possessing clear and convincing evidence seems the correct threshold for assertion, flat out belief, relying on belief in action, warranting knowledge attribution, and so on. And plausibly the kinds of state actions legitimated by a finding to the clear and convincing evidence standard are the kinds of state actions that require knowledge of the litigated claims. This includes, for example, right-to-die hearings, child custody, paternity disputes, and commitment to mental institutions.[15]

It is worth emphasising that the overall framework articulated in this essay, which describes the three legal standards of proof, can be assessed independently from the auxiliary claims that (i.) knowledge requires ruling out relevant alternatives, and (ii.) the tier of error possibilities relevant to the 'clear and convincing evidence' standard matches the error possibilities relevant to knowledge ascriptions. One might adopt the overall framework for interpreting the three legal standards and eschew the relevant alternatives account of knowledge. And one could endorse a relevant alternatives account of knowledge and a relevant alternatives account of the clear and convincing evidence standard, for example, but posit that knowledge-relevant alternatives—unlike those relevant to the legal standard—vary wildly with conversational context.

Beyond reasonable doubt is a more demanding standard. As its name suggests, it requires addressing all reasonable error possibilities. This is a more expansive set than the alternatives relevant to clear and convincing evidence.

---

[14] The *OED* supplies two definitions for 'preponderant': (i) Greater or superior in influence, power, or importance; predominant. (ii) Greater in weight, heavier. For 'preponderance' it further supplies a third meaning: (iii) Superiority in number or amount; an abundance, a great number, a large quantity, a majority. On this proposal, the first definition informs the legal standard and advises against a quantitative interpretation.

[15] The standard is often interpreted as around 70 or 75%. See United States v. Fatico, 458 F. Supp. 388 (1978), McCauliff (1982), Clermont (1986–87), Sand and Rose (2003), and Moss (2018a, 2018b). Clermont (1986–87: 1119) writes of 'the intermediate standard or standards, often grouped under the banner of clear and convincing evidence and roughly translated as much-more-likely-than-not. Judicial formulations include "clear, cogent, and convincing," "clear, satisfactory, and convincing," "clear, precise, and indubitable," clear and irresistible," and "convincing beyond reasonable controversy."'

Some epistemic standards are even more demanding than beyond reasonable doubt. The threshold 'beyond all possible doubt', for instance, is available. The sphere of error possibilities that must be eliminated to satisfy this standard is extremely large, perhaps encompassing *all* error possibilities. Given that further sub-possibilities can usually be described, very few claims can be satisfied to this extremal standard. For almost every claim there are some error possibilities not eliminated by the evidence. This underdetermination fuels the skeptical challenge.

The overall framework, then, adapts the relevant alternatives structure for knowledge and instead proposes a relevant alternatives condition on legal standards of proof. This yields:

- General: Claim p is established to a legal standard L only if the evidence adduced rules out the L-relevant error possibilities.
- Preponderance: Claim p is established to a preponderance of the evidence only if the evidence adduced rules out preponderant error possibilities.
- Intermediate: Claim p is established to the 'clear and convincing evidence' standard only if the evidence adduced rules out the error possibilities relevant to the clear and convincing evidence standard. (Auxiliary: These error possibilities are the knowledge-relevant ones.)
- Reasonable doubt: Claim p is established beyond reasonable doubt only if the evidence adduced rules out the reasonable error possibilities.

Plausibly standards like 'reasonable suspicion' and 'probable cause', which govern police conduct, have a different epistemic structure. They are not verdictive, and so to satisfy the standards, evidence need not eliminate classes of error possibilities. This helps explain, for instance, why police can permissibly detain and arrest more individuals for a crime than they know can be guilty of that particular crime. The police must meet epistemic thresholds before acting, but those thresholds plausibly are discontinuous with the three verdictive legal standards described here.

To better illuminate this framework, consider an example. Return to the litigated claim:

R: The defendant made an agreement with the complainant.

As noted above, the complainant alleges R; the defendant denies. There are many error possibilities—many ways the claim might be false. Some of these sub-possibilities are mundane and others farfetched. Suppose the case is heard in civil court and so is governed by the preponderance standard. To establish R to this standard, the complainant need not address farfetched error possibilities, such as that the world is only five minutes old or the city cabals against her. She also need not address many of the less farfetched error possibilities. The evidence must address all the preponderant error possibilities. These are the sub-alternatives that are most significant, most predominant. They are the most usual ways that claims like R are false; the ones a judge would most expect to see in her courtroom. They are the error possibilities a virtuous inquirer would think most important to investigate and that seem most weighty as potential doubts about R.

There are many such ordinary possibilities: She misunderstood her communications with the defendant. She was overly optimistic or insufficiently attentive during discussions. She mistook positive reactions from the defendant as a binding agreement. He said 'I would love to build your conservatory at that price'; he meant 'So let's see if we can agree on detailed terms'. She took his statement to constitute an agreement. She misremembers the outcome of the negotiation. She has confused him with someone else. Her inexperienced assistant naïvely assumed an agreement was reached, and she relied on this assistant. Self-servingly she is intentionally overstating what was said. She has already acted on the assumption of a forthcoming agreement and consequently confronts

financial problems that would be solved if he now agrees to the proposed contract. This looming financial problem prods her towards usual human prevarication. These are ordinary, typical ways that claims like R are false. These are preponderant on the mind of the judge. If we were inquiring into whether R, these are the first alternatives we would think to address. To establish the claim to the preponderance standard, her evidence must eliminate these error possibilities.[16]

Suppose the complainant produces documentation—a written contract with the defendant's signature. This evidence addresses those preponderant error possibilities. In most cases, this additional evidence establishes R to the preponderance standard. (The example aims to illustrate and motivate the 'three concentric thresholds' feature of the relevant alternatives model of standards of proof. Disputes concerning precisely how evidentially demanding the three legal standards are can be set aside for now.)

Note the evidence can address relevant alternatives without anyone thinking explicitly in terms of error possibilities or remoteness. As with the zebra example above, we appeal to evidence concerning what *is* the case; we need not think of ourselves as addressing possibilities about what is not the case. Ruling out error possibilities is normal and automatic: 'Look, we had an agreement. See this signature!'

Suppose now that R must be established to a more demanding standard of proof, the clear and convincing evidence standard. Perhaps the agreement concerns child custody, for example. The evidence adduced above might be insufficient to satisfy this higher standard. It fails to eliminate various error possibilities—various sub-alternatives of not R. It does not rule out that she faked the defendant's signature, tricked him into signing something by deviously splicing two documents, or confused the defendant with someone else of the same name. Some of these error possibilities are properly disregarded for the lower preponderance standard, but are relevant to the clear and convincing evidence standard. To meet this more demanding standard she must address these further error possibilities. She must thus produce further evidence.

Suppose she has additional evidence. The defendant's secretary testifies that on the morning of the contract's date the defendant expressed his intention to sign the contract. This evidence does not fully eliminate all remaining error possibilities—almost no evidence could—but it addresses several of the central ones. It is still possible the complainant faked the defendant's signature. But these remaining sub-possibilities require, for instance, that the defendant changed his mind or his secretary lies. Plausibly these are marginal enough sub-sub-possibilities to be disregarded for the clear and convincing evidence standard. (Bear in mind they are sub-sub-sub possibilities. On this branch of sub-alternatives, there is no agreement, the complainant is lying to frame the defendant, she has faked his signature on a contract, and either the defendant misrepresented his plans to his secretary or the secretary commits perjury against their boss's interests.)

Thus, with the introduction of the secretary's testimony, the remaining uneliminated error possibilities are remote enough to disregard for this standard. The evidence establishes R to the clear and convincing evidence standard.

Finally suppose the putative agreement concerns a criminal matter, and so claim R must be established to the beyond reasonable doubt standard. This standard is higher. To satisfy beyond reasonable doubt, all reasonable error possibilities must be eliminated by the evidence. The 'relevant

[16] Note that in no sense must she eliminate *most* of the error possibilities or *most* of the preponderant error possibilities. We ought not quantify over possibilities; it is hard to understand what this could mean. Instead she must eliminate all preponderant error possibilities (or the most preponderant, where 'most' is interpreted qualitatively, as discussed above). Admitting comparisons does not commit us to quantification. Some beaches are more beautiful than others; we need not quantify the beaches or the beauty.

alternatives' interpretation of the reasonable doubt standard is supported by explicit statements in legal education resources and judges' comments, such as,[17]

> [P]roof beyond a reasonable doubt is such proof as precludes every reasonable hypothesis except that which it tends to support, and is proof which is wholly consistent with the guilt of the accused, and inconsistent with any other rational conclusion.

If the investigation is governed by the reasonable doubt standard, the above uneliminated error possibilities cannot be disregarded. Perhaps the defendant changed his mind before committing, causing the complainant to fake his signature, for example, or perhaps the defendant never expressed an intention to sign, but the secretary and complainant conspire for financial gain. Plausibly these possibilities are a source of reasonable doubt; they are uneliminated error possibilities that fall within the broader sphere circumscribed by the reasonable doubt standard. (Again, qualms about precisely how evidentially demanding the legal standards are can be set aside, since the example is to illustrate the structure of the framework.)

Suppose the document adduced by the complainant includes a notary stamp. The notary has, furthermore, submitted an affidavit attesting that she retains records of the signing event. This evidence addresses the remaining reasonable error possibilities. Claim R is, given this further evidence, established beyond reasonable doubt. Some further possible doubts remain uneliminated. Perhaps the notary conspires with the complainant and the secretary, for instance. But such doubts are not reasonable. They are fanciful and can be disregarded, even for this demanding standard.[18] As Thomas Starkie (1933: 514) writes,

> To acquit upon light, trivial or fanciful suppositions, and remote conjecture, is a virtual violation of the juror's oath, and an offence of great magnitude against the interests of society… On the other hand, a juror ought not to condemn unless the evidence excludes from his mind all reasonable doubt as to the guilt of the accused.

The second sentence expresses something akin to the relevant alternatives account of the reasonable doubt standard. The first sentence articulates a limit on what error possibilities qualify as relevant.

4. Remoteness

This model is schematic and leaves open various choice points. The framework posits that error possibilities are increasingly remote. Those that are sufficiently typical must be taken seriously by the court. If they cannot be addressed by the evidence, the standard of proof is not met. An error possibility is addressed if every sub-possibility is either eliminated by the evidence or lies beyond the threshold of normality and so need not be taken seriously by the court. There are three cut-off thresholds, corresponding to the increasingly demanding standards of proof.

The model is consistent with various conceptions of evidence, possessing evidence, and eliminating possibilities. Two key features to determine are (i.) What makes some error possibilities remote and others close by? (ii.) What determines the location of the three thresholds of relevance? Both questions ask why some error possibilities are disregardable, whilst others must be addressed in order to find the defendant culpable.

We have intuitive judgements, at least roughly, concerning which error possibilities must be addressed. We know we need not address skeptical or paranoid doubts in order to convict, for

---

[17] *Corpus Juris Secundum*, vol. 23, *Criminal Law*, §1502. For further examples and instructive discussion, see Ho (2008: 156–161).

[18] In corrupt societies such doubts might be reasonable. As I describe below, one cost of official corruption is rendering reasonable many more error possibilities, thereby making legal standards harder to satisfy.

example. And if three independent witnesses claim p, then, barring significant counterevidence, p is established at least to a preponderance of the evidence. If we lack access to at least central examples of error possibilities located either side of these thresholds, this threatens my proposal. But it threatens the entire court system too.

Many sources inform this question, including legal practice, legal training, explicit laws, court records, precedent, convention, and legal practitioners' reflective judgements about cases. Criticisms of legal judgements are revealing. Theorists criticise a verdict because the factfinder treated as unreasonable an error possibility that on reflection seems reasonable, or vice versa. For civil suits, a factfinder and commenter can disagree about what error possibilities are predominant.[19]

Epistemologists might additionally theorise about what underwrites these intuitive facts.[20] Perhaps remoteness is determined by what is normal or typical in the actual world, or by which alternatives most resemble the actual world.[21] Perhaps which error possibilities we must eliminate is determined by human psychology—how defendants, witnesses, and factfinders act—or by legal track records, such as what is frequently true when litigated claims are false. Perhaps remoteness is partly determined by what virtuous inquirers would rule out, what is important to rule out, or what is easy to rule out.[22]

The thresholds might be determined by norms of action. Norms describe the epistemic requirements for legitimating state actions such as assertion of guilt, criminal punishment, holding responsible, imposing fines, and determining life paths such as involuntary institutional confinement, child custody, and so on.

Crucially, which error possibilities are nearby or farfetched is determined, at least in large part, by what society is actually like and what tends to occur. The resulting view is thus externalist—whether error possibilities are relevant to a legal threshold is not determined by factors wholly accessible to the factfinder. And these possibilities can change over time. If defendants and witnesses tend to fake X, then faking X is a nearby possibility. If police officers plant guns and drugs on unarmed citizens, then this possibility is a source of reasonable doubt.

If we adopt the Lewisian claim that the actual is always a relevant alternative, this entails that no false judgement can be established to a legal standard. It is an open question whether this consequence is attractive. I will not evaluate its attractiveness here, but note that several competing accounts of the epistemology of legal proof—such as safety, sensitivity, causation, and most knowledge-based accounts—are committed to this consequence.[23] A relevant alternatives account of legal proof can avoid this consequence by rejecting Lewis's claim.

A relevant alternatives account of legal proof should reject the Lewisian claim that an alternative is relevant if under consideration. There are at least two reasons. Firstly, Lewis's account of knowledge entails that knowledge attributions are fickle—they can be true one minute and false the

---

[19] We criticise proceedings for other reasons too, of course, such as disagreement about whether particular evidence is admissible.

[20] See Stine (1976), Lewis (1996), Hawthorne (2004), Ho (2008), Dougherty and Rysiew (2009), Fantl and McGrath (2009), McKinnon (2013), Amaya (2015: 525ff), Moss (2018a, 2018b), and Gardiner (ms-c).

[21] Theorists might combine the proposed framework with an account of remoteness from Smith (2010; 2016) or Pritchard (2005; 2018). Gardiner (2017b, forthcoming, ms-b, ms-c) articulates objections to a Pritchardian view on which remoteness is determined solely by similarity to the actual world.

[22] Perhaps it is atypical for a complainant to have mistaken the defendant for the person she in fact made an agreement with, for example, yet such error possibilities are predominant because easily addressed. If so, relative cost of inquiry affects relative remoteness. Some error possibilities are more morally significant than others. Misidentifying the defendant is usually a costlier error than miscategorising the crime, for example. Moss (2018a, 2018b) and Gardiner (ms-c) discuss whether differentiated error costs affect relative remoteness of possibility.

[23] See Blome-Tillman (2015) and Gardiner (2018). Duff et al. (2007), Ho (2008), Moss (2018b), and Gardiner (forthcoming) discuss the factivity of proof beyond reasonable doubt.

next. But legal proofs are stable; they cannot dissolve and reappear because a creative juror has conjured or forgotten an error possibility.[24] Secondly, what jurors consider is over- and under-inclusive. Prejudiced jurors might fail to consider alternatives in which a Black defendant is innocent, for instance, or forget crucial exculpatory testimony. Poor juror thinking does not affect which error possibilities ought to be addressed. Relevance might depend on what jurors ought to consider, but it cannot depend on what they do consider.[25]

The framework as articulated is ecumenical about whether practical stakes can affect the location of the thresholds from case to case. Practical features might affect why the criminal standard is higher than the intermediary and preponderance standards. But this observation is compatible with each threshold being fixed. Theorists can adopt this framework and endorse either fixed or stakes-sensitive thresholds for the legal standards.[26]

5. Virtues

In this section I articulate virtues of the relevant alternatives framework for interpreting legal standards of proof. I first sketch problems afflicting rival views that the relevant alternatives framework avoids. I then articulate aspects of legal normativity and practice that the framework illuminates.

The relevant alternatives framework explains the inadequacy of bare statistical evidence for legal proof. Consider the following vignette, adapted from Cohen (1977).

> Gatecrasher: A rodeo sells admittance, but does not issue tickets. One day the gate is left open and, taking advantage of the lack of ticketing system, many people gatecrash. The managers realise only 40 seats were sold, but 100 people attended. Hoping to regain some of their financial loss, the rodeo organisers file small civil claims against arbitrarily selected attendees.

The bare statistical evidence renders it numerically likely that each defendant gatecrashed. Yet the evidence seems inadequate to secure a finding, even to the preponderance standard. This is hard to explain if legal standards of proof are a numerical likelihood threshold. The relevant alternatives framework explains why bare statistical evidence cannot satisfy the legal standards. The non-remote error possibilities in which defendants are not culpable cannot be properly ignored, and the evidence does not address these possibilities.

There are various approaches available for establishing why the Gatecrasher error possibility is not remote. The details depend on how remoteness is determined, but—as far as I know—every relevant alternatives account of knowledge includes resources to explain Lottie's lack of knowledge. (Lottie, introduced in section two, reasons statistically to conclude her lottery ticket lost.) Unless the explanation relies on fleeting features of context, these resources also explain the inadequacy of bare statistical evidence for legal proof. A relevant alternatives account of legal proof can draw on additional explanatory resources, furthermore, including the aims of *legal* inquiries, the significance of affirmative verdicts in legal proceedings, and the special relevance of *morally significant* error possibilities.[27]

---

[24] Lewis (1996: 560). See also Amaya (2015: esp. 517) and the imaginative juror from Barber and Gordon (1976: 76), quoted in Ho (2008: 153).

[25] A similar problem afflicts Moss's (2018b: 221) 'norm of consideration' explanation of an epistemic fault of profiling. Briefly: The norm notes 'you morally should keep in mind the possibility that [a person] might be an exception to statistical generalizations'. Here is the problem: If someone follows the norm, the error possibility is 'in mind', and thus relevant. But if they fail, the unconsidered error possibility may remain irrelevant, and the belief can qualify as knowledge. See Moss (2018a: 190–2) and Gardiner (ms-c) for discussion.

[26] See Ho (2008), Amaya (2015), Gardiner (ms-c).

[27] Lewis (1996), Ho (2008), Amaya (2015), Moss (2018b). Gardiner (ms-c) discusses whether particular error possibilities can be relevant in virtue of being morally differentiated. See also Moss (2018a; 2018b), Bolinger (forthcoming).

Various competing proposals have been developed to explain the inadequacy of bare statistical evidence in cases like Gatecrasher. Many proposals endorse the underlying foundation that standards of proof are numerical likelihood thresholds—typically either a measure of credence or evidential probability—but augment that foundation with further necessary conditions on legal proof. These include, for example, that the verdict is safe or sensitive, the evidence is causally related to, or normically supports, culpability, or the evidence is knowledge conducive. Although I lack space to evaluate them here, these accounts face problems explaining the inadequacy of bare statistical evidence when applied to a fuller range of vignettes.[28] But even if such accounts can explain the inadequacy of bare statistical evidence for legal proof in a variety of cases, they confront other problems. By endorsing the quantified likelihood foundation, they inherit problems generated by that foundation. The relevant alternatives framework eschews this foundation, and so avoids these problems.[29]

One such problem is the challenge of determining the initial likelihood of culpability at the trial's outset. If legal evidence is amenable to mathematical treatment, such as Bayesian updating, then what is the antecedent likelihood of culpability before evidence is adduced? Perhaps it is 50%; but since this far exceeds everyday levels of suspicion, this plausibly conflicts with the presumption of innocence. And what if the trial evaluates more than one count or a three-way decision between crime X, crime Y, and not guilty? Perhaps prior probability is sensitive to base rates of guilt in the defendant population—but these might be difficult or impossible to discern, and may vary by crime, time, jurisdiction, and social demographic.[30] Plausibly such features should not affect how near the inculpatory case is to satisfying the standard of proof from the outset. Refusing to provide an antecedent likelihood of culpability at the outset of the trial undermines the mathematical approach to legal standards of proof. But any given number seems unjust, arbitrary, naïve, or all of the above.

A second worry is that any particular numerical threshold, such as 95%, is arbitrary. A third worry concerns how to aggregate judgements of likelihood within a group of factfinders. Some proceedings have a three-magistrate bench, others a twelve-member jury. To satisfy a particular numerical likelihood threshold, must each individual's judgement reach that confidence level, or does the threshold apply to the group's aggregated judgement? If the latter, how should the judgements be aggregated? This question is fraught.[31]

A fourth problem afflicting quantitative probabilification accounts of legal proof is that in ordinary cases, people do not have fine-grained judgements about quantified statistical likelihood. Individuals cannot determine whether evidence renders guilt 89% likely compared to 96% likely, for example. Unless examples concern coin tosses, dice rolls, lotteries, and so on, ordinary individuals do not use quantitative probability estimates. (Even in those cases, they are error prone.) The problem is not mere inaccuracy about such judgements; the problem is that plausibly individuals do not normally

---

[28] Pritchard (2018), Smith (2016), Thomson (1996). Note Blome-Tillman's (2017) and Littlejohn's (2018) knowledge-based accounts of legal proof focus on the likelihood of knowledge of culpability, not the likelihood of culpability itself, thereby avoiding some of these problems. I critically engage with these accounts elsewhere (Gardiner 2018; forthcoming; ms-b), but lack space here. Gardiner (ms-b) argues, for example, that even though Lottie's lottery belief could easily be wrong, some beliefs based on bare statistical evidence are modally secure. I thereby dispute the safety-based explanation of the inadequacy of bare statistical evidence for legal proof. See also Blome-Tillman (2015) and Redmayne (2008: 301–2).

[29] For problems plaguing quantifiable balance approaches to legal proof see, for example, Tribe (1971), Cohen (1977), Ho (2008; 2015), Haack (2014), Amaya (2015), Nance (2016), Smith (2016), Pardo (2019). For overviews, see Ho (2013) and Gardiner (2019). Related problems plague quantifiable balance accounts of justified belief. See Nelkin (2000), Achinstein (2003), Buchak (2014), Staffel (2016), Jackson (2018).

[30] Perhaps white-collar crime is unlikely to reach the courtroom, for example, unless the defendant is guilty. King (ms) argues these base rates are unknowable.

[31] See, for example, Fitelson and Jehle (2009), List and Pettit (2011), Goldman and Whitcomb (2011: chs 10 and 11) Russell, Hawthorne, and Buchak (2015), and Easwaran, Fenton-Glynn, Hitchcock, and Velasco (2016).

make quantified balance judgements at all. It is not how we normally reason. The relevant alternatives framework more plausibly models how thinkers reach verdicts.

The relevant alternatives model of legal standards of proof avoids these problems by eschewing the numerical framework. In addition to avoiding these problems, the proposed framework enjoys various virtues. It illuminates features of legal normativity and practice, some of which are rendered mysterious by rival accounts.

Firstly, it explains and systematises the epistemic error of large classes of poor legal judgements. If a judgement treats as beyond the threshold, and so ignorable, an error possibility that is within the threshold, that judgement is criticisable. Suppose there is an error possibility—a sub-alternative in which the defendant is innocent—that is consistent with the evidence adduced. But it includes that four police officers are all lying under oath and one planted drugs in the defendant's car. The jurors might treat this error possibility as too outlandish to be relevant, and so disregard it. They find guilt beyond reasonable doubt. If such police behaviour is in fact not remote—if the source of doubt is reasonable—we can criticise the decision.

Conversely a decision can be criticised if the factfinders treat a relatively remote error possibility as nearby, and so relevant to the legal standard. Suppose three women each accuse a wealthy celebrity of sexual violence, and one of these women sues for damages. At court the three testimonies describe similar conduct and, furthermore, they testify they had not met before talking to investigators. The jurors take seriously the error possibilities—the possibilities in which the celebrity is innocent. They consider the possibility that the women are secretly friends who concocted their stories for financial gain. The jurors decide this error possibility is preponderant and uneliminated. They accordingly find in favour of the defendant. In this case the factfinder has mistaken a remote possibility for a preponderant one.

Such examples exemplify how the framework highlights and taxonomises sources of systematic epistemic injustice in legal proceedings. If police malfeasance and perjury is consistently treated as a remote possibility, but is not remote, this is systematic epistemic legal injustice. If women's lying about sexual assault is remote, abnormal, and rare, yet is deemed typical and commonplace and accordingly treated as a predominant error possibility, this is systematic epistemic legal injustice.[32]

The relevant alternatives framework highlights an epistemically pernicious feature of practices of evidence tampering and police perjury. Perhaps the officers tell themselves, 'we only do this infrequently, and only when needed to secure convictions for violent criminals.' But police doing this infects the epistemology of other trials. The practice makes evidence tampering a relevant possibility. If evidence tampering is a relevant alternative, it becomes substantially harder—perhaps impossible—to prove any crime beyond reasonable doubt. Tampered evidence is a difficult error possibility to eliminate; once relevant it will remain an uneliminated relevant alternative. To secure convictions, evidence-tampering possibilities must remain farfetched. For this to happen, they must in fact *be* farfetched.

The framework highlights a potential epistemic pitfall to which reasoners are vulnerable. On the account I propose, merely considering an error possibility does not make that possibility more relevant or more reasonable. The remoteness of an error possibility is fixed. The thresholds of the three legal standards are also not affected by what the factfinder dwells on. The evidence can address a relevant (that is, sufficiently nearby) error possibility by either being inconsistent with the possibility or—much more commonly—being inconsistent with many sub-alternatives of that error possibility, and leaving uneliminated only those sub-alternatives that are too remote to be relevant. But the

[32] Gardiner (ms-d) argues false accusations of acquaintance rape are an unusual occurrence and investigates whether, given they are typically true, rape accusations alone can satisfy the 'preponderance' standard. Gardiner (ms-a) applies the relative alternatives framework to the epistemology of rape accusations.

evidence, by doing this, can thereby draw attention to these remote error possibilities. The factfinder might never have even thought of the possibility unless prompted by the evidence. And attention is thereby drawn to the fact that her evidence does not eliminate those possibilities. The very process of eliminating error possibilities can thereby make remote error possibilities seem important. This is because factfinders thereby think about them as uneliminated and confront the fact that their evidence is consistent with error possibilities. This itself does not make the possibilities relevant. But it can make the possibilities *feel* relevant. As better and better evidence stacks up against the wealthy celebrity, or increasingly indicates that the earlier defendant did have an agreement with the complainant, this can have the unintended consequence of drawing attention to the convoluted sub-sub-sub alternatives in which the litigated claim is false. Factfinders can thereby mistakenly think of such errors as preponderant or significant.

A related phenomenon is that when evidence addresses a sub-alternative, that evidence can make a subset of that sub-alternative more probable. But that uneliminated error possibility nonetheless remains beyond the relevant threshold, and so too remote to take seriously. Suppose ten years prior to the lawsuit—predating the celebrity's fame—one of the women told her therapist of the attack, including naming him. This evidence rules out many error possibilities, such as those in which she recently fabricated the story. But the evidence makes some (remote, outlandish, unreasonable) error possibilities *more* likely, such as that she fabricated the story years ago because of a lifelong obsession with the defendant. This error possibility is, in almost every case, itself too outlandish to require addressing. Factfinders would not have even thought of it prior to hearing the inculpatory evidence concerning therapy. But it is nonetheless made more probable by the evidence. The corresponding epistemic pitfall is mistaking the fact that the probability of the remote error possibility has increased, given the evidence, for the claim that the error possibility is now relevant and ought be taken seriously.[33]

Appeal to relevant alternatives illuminates real-life legal verdicts that are otherwise puzzling. The inadequacy of bare statistical evidence for affirmative verdicts is often illuminated using a traffic accident vignette.[34] The accident was caused by an unidentified hit-and-run bus. Since most buses in town are operated by Red Bus company, the plaintiff sues Red Bus. The plaintiff reasons that probably, given the base rates, the culpable vehicle was operated by Red Bus, so she ought to win. The example is loosely based on a real-life lawsuit, *Smith v. Rapid Transit*, but the real-life case did not involve any statistics or base rates.

Late at night on February 6th 1941 Betty Smith was driving along Main Street in Winthrop, Massachusetts, when an unidentified bus forced her to swerve into a parked car. She later discovered that only one company, Rapid Transit, Inc., was franchised along that street, and the route ran frequently throughout the night. A second company operated elsewhere in town. With this inculpatory evidence, Smith sued Rapid Transit.

The judge ruled in favour of Rapid Transit. He reasoned that there was no 'direct evidence' that the offending bus was operated by Rapid Transit. The verdict says,[35]

---

[33] Indeed perceiving how new inculpatory evidence—telling the therapist—can be consistent with the defendant's innocence—such as if the plaintiff harbours a lifelong obsession—can produce an 'ah-ha' sensation that resembles realising the truth. The evidence 'fits' the new error possibility. Thinking of such possibilities, and seeing how they are consistent with the evidence, is a kind of intellectual accomplishment. The psychological feeling accompanying thinking up a farfetched error possibility that is consistent with severely damning evidence can, in other words, delusively suggest the defendant's innocence. I am grateful to Julien Dutant, Bruce Chapman, Catherine Elgin, Jon Garthoff, and Hilary Kornblith for helpful discussions on these topics.

[34] Tribe (1971), Thomson (1986), Buchak (2014).

[35] Appellate verdict, Smith v. Rapid Transit, Inc. 317 Mas. 469, 58 N.E.2d 754 (1945).

> While the defendant had the sole franchise for operating a bus line on Main Street, Winthrop, this did not preclude private or chartered buses from using this street; the bus in question could very well have been operated by someone other than the defendant… The most that can be said of the evidence in the instant case is that perhaps the mathematical chances somewhat favor the proposition that a bus of the defendant caused the accident. This was not enough.

Many existing explanations of the inadequacy of bare statistical evidence for legal proof do not impugn the non-statistical evidence adduced in *Smith v. Rapid Transit*. If an affirmative verdict must be safe, for example, as Duncan Pritchard (2018) claims, this might explain the inadequacy of the Gatecrasher evidence, but Betty Smith's evidence *does* satisfy the safety condition. In all nearby worlds where Smith possesses the evidence, her belief is true.[36] The relevant alternatives framework illuminates the judge's decision. He judged that the error possibility 'the bus in question [was] operated by someone other than the defendant' was preponderant. And Smith's evidence, as the judge explicitly notes, does nothing to address this error possibility, because the 'sole franchise' does not 'preclude private or chartered buses'. Those who debate whether the ruling was correct might focus on the statistical probability of the bus not being a Rapid Transit bus. They should instead focus on whether this error possibility is significant enough to qualify as preponderant, as the judge held, or remote enough to properly ignore. Statistical likelihoods inform, but do not resolve, this question.[37]

Finally, the relevant alternatives framework can explain the appeal of competing approaches. Duncan Pritchard (2018) argues that affirmative verdicts must be safe. A verdict is safe iff not easily could it have been erroneous. Pritchard develops this idea using a possible worlds framework. Roughly, a verdict is safe iff in all nearby worlds where the verdict is formed using the same method as in the actual world, the verdict is true. This condition reflects the crucial idea that for judgements of legal culpability, it is not sufficient that the factfinder is statistically likely to be right. As Pritchard (2018: 117) writes, 'what is required for a conviction is evidence such that, given that evidence, it cannot be an easy possibility that the defendant is wrongfully convicted.'

But notice that the relevant alternatives framework does justice to this idea. On the proposed framework, the most predominant errors must be ruled out. The most common and ordinary ways the litigated claim could be false must be addressed. The only remaining error possibilities are distant and strange, such as where the entire city conspires against the defendant. In effect, the relevant alternatives framework demands that erroneous verdicts of culpability could not easily happen; they happen only in remote circumstances.[38]

David Enoch, Levi Spectre, and Talia Fisher (2012) posit that to satisfy legal standards of proof, the evidence adduced must be sensitive to culpability; if the defendant were innocent, the

---

[36] I develop this argument in Gardiner (forthcoming). *Smith v. Rapid Transit* does not involve any statistical evidence and so differs significantly from statistical evidence examples like Gatecrasher and Red Bus. Gardiner (ms-b) argues that modal accounts, like Pritchard's safety account, also cannot explain the inadequacy of bare statistical evidence exhibited in fictional vignettes like Gatecrasher and Red Bus.

[37] Gardiner (ms-c) describes how the relevant alternatives framework explains epistemic flaws in many instances of profiling. The epistemic features of profiling—understood as forming opinions about an individual on the basis of features believed about social group or characteristics—bear some similarity to the epistemic inadequacy of bare statistical judgement in cases like Gatecrasher but, I argue, have underappreciated differences.

[38] The overall framework I motivate, which rivals the widely-held quantifiable balance framework, is consistent with various substantive accounts of remoteness. A theorist might adopt the proposed framework, for example, and posit that remoteness is determined solely by similarity to the actual world. The resulting view would resemble a Pritchardian safety account, augmented with three thresholds of relevance, corresponding to the three legal standards. Elsewhere I argue this Pritchardian approach will fail, however, because (i.) some beliefs based on bare statistical evidence are safe, (ii.) a remoteness ordering based solely on similarity to the actual world cannot explain why misleading evidence renders error possibilities relevant, and (iii.) the value of safety is swamped by the value of truth. I lack space to do justice to these claims here. See Gardiner (2017b; forthcoming; ms-b).

evidence would have been different, and so the factfinder would not have believed the defendant were culpable. This too captures an important aspect of legal justice. We should find people culpable only if their culpability makes a difference to the judgement. When ruling against the defendant, the state should be able to assert that if the accused did not the commit the crime, he would not have been convicted. It is a plausible principle of justice and good reasoning that a person should not be found culpable unless the conviction is sensitive to the transgression.[39]

The relevant alternatives framework vindicates this idea. Sensitive evidence is valuable, on this framework, because such evidence is how relevant error possibilities are addressed. The signed, notarised document is sensitive to litigated claim R, for example. If there were no agreement, then—within relevant limits—the document would not have obtained. The evidence, in virtue of its sensitivity, eliminates error possibilities.

Judith Jarvis Thomson (1986) argues that for legal proof, evidence must be causally related to culpability. This is because, she argues, legal verdicts require guarantees. Only causally related evidence—not bare statistical evidence—can underwrite a guarantee. This proposal has appeal. Plausibly legal consequences—even those regulated by the relatively low preponderance standard—rely on *establishing* claims. They treat the claims as reasons for action.[40] For this to be legitimate, some security is required; some guarantee. Judges' opinions refer to the preponderance standard demanding 'actual belief', 'proof', 'establishing as a fact', and 'knowledge'.[41] Such ideas suggest that finding to the preponderance standard is considered a kind of guarantee. Mere numerical chance, especially as low as 50%, is insufficient.

The relevant alternatives framework vindicates a sense in which a legal verdict of culpability, even to the mere preponderance standard, constitutes a (weak) guarantee. The pressing error possibilities are addressed. Evidence eliminates any normal ways in which the defendant is innocent. If the verdict is incorrect, something farfetched and unexpected has occurred.[42] The factfinder has considered the central error possibilities, and the evidence has ruled them out. This provides a kind of guarantee—one not supplied if legal standards demand simply a quantifiable likelihood of guilt.[43]

6. Conclusion

The relevant alternatives framework for legal standards of proof is promising. It provides resources to improve upon the commonly-held 'quantifiable balance' conception in at least two central ways. Firstly, by providing a non-numerical foundation it avoids problems generated by numerical approaches. Secondly, by focusing on error possibilities, the proposed framework highlights the epistemic significance of missing evidence. Even if the evidence considered renders culpability likely, it cannot underwrite an affirmative verdict if that evidence leaves relevant error possibilities unaddressed. High evidential balance is not enough—sometimes more evidence is needed. And, as the Gatecrasher and real-life *Rapid Transit* cases illustrate, the relevant alternatives framework can explain when and why. The relevant alternatives framework thereby improves upon both the 'quantifiable' and 'balance' aspects of the quantifiable balance conception of legal standards of proof.

---

39 Enoch, Spectre, and Fisher (2012) and Enoch and Fisher (2013) develop an incentive-based account of the legal value of sensitivity.

40 Littlejohn (2018).

41 See, for example, *Smith v. Rapid Transit*; *Day v. Boston & Maine Railroad*. Focusing on eliminating error possibilities, rather than merely having evidential balance on your side, also helps emphasise the asymmetry of the proof *burden*. It must be established by one side; it is not simply a symmetrical matter of whether claims are likelier true or false.

42 See Smith (2016).

43 Finally, theorists who endorse a relevant alternatives framework for knowledge can thereby explain, or explain away, the sense that legal proof requires knowledge. The theorist can argue that legal proof and knowledge are characterised by the same foundational epistemic feature—namely, they require eliminating relevant alternatives.

Acknowledgements
I am grateful to Dominic Alford-Duguid, Grace Boey, Renee Bolinger, Rebecca Brown, Bruce Chapman, Antony Duff, Marcello Di Bello, Julien Dutant, Catherine Elgin, Branden Fitelson, Jamie Fritz, Mikkel Gerken, Doug Garthoff, Jon Garthoff, Zachary Hoskins, Doug Husak, Hilary Kornblith, David Lagnado, Youngjae Lee, Clayton Littlejohn, Sebastian Liu, Silvia Milano, Beau Madison Mount, Jessie Munton, Federico Picinali, Anne Ploin, Maura Priest, Duncan Pritchard, Paul Roberts, Bernhard Salow, Paul Silva, Martin Smith, Alec Walen, Aness Webster, William Wells, Tim Williamson, Alex Worsnip, and three anonymous referees for invaluable discussions and feedback. Thanks also to participants at the Epistemology and Criminal Evidence workshop at Nottingham University School of Law, the Between Ethics and Belief conference at Cologne University, and the Moral Encroachment Reading Group at Oxford University.

Bibliography
Achinstein, Peter (2003) *The Book of Evidence* OUP.
Amaya, Amalia (2015) *The Tapestry of Reason* Hart Publishing.
Austin, John (1946) 'Other Minds' *Proceedings of the Aristotelian Society*, *Supplementary Volumes* 20:148–187.
Barber, Dulan and Giles Gordon (eds.) (1976) *Members of the Jury* Wildwood House.
Basu, Rima (forthcoming) 'The Wrongs of Racist Beliefs' *Philosophical Studies*.
Blome-Tillmann, Michael (2015) 'Sensitivity, Causality, and Statistical Evidence in Courts of Law' *Thought* 4(2):102–112.
________ (2017) '"More Likely Than Not" Knowledge First and the Role of Bare Statistical Evidence in Courts of Law.' *Knowledge First* Adam Carter, Emma Gordon, and Benjamin Jarvis (eds.) OUP, 278–292.
Bolinger, Renee (forthcoming) 'The Rational Impermissibility of Accepting (Some) Racial Generalizations' *Synthese*.
BonJour, Laurence (2010) 'The Myth of Knowledge' *Philosophical Perspectives* 24(1):57–83.
Bradley, Darren (2014) 'A Relevant Alternatives Solution to the Bootstrapping and Self-Knowledge Problems' *Journal of Philosophy* 111(7):379–393.
Buchak, Lara (2014) 'Belief, Credence, and Norms' *Philosophical Studies* 169(2):285–311.
Clermont, Kevin (1986–1987) 'Procedure's Magical Number Three: Psychological Bases for Standards of Decision' *Cornell Law Review* 72:1115–56
Cohen, Jonathan (1977) *The Probable and the Provable* Oxford University Press.
Cohen, Stewart (1986) 'Knowledge and Context' *Journal of Philosophy* 83(10):574–583.
Di Bello, Marcello (2013) *Statistics and Probability in Criminal Trials: The Good, the Bad and the Ugly* doctoral dissertation, Stanford University.
Dougherty, Trent and Patrick Rysiew (2009) 'Fallibilism, Epistemic Possibility, and Concessive Knowledge Attributions' *Philosophy and Phenomenological Research* 78(1):123–132.
Dretske, Fred (1970) 'Epistemic Operators' *Journal of Philosophy* 67(24): 1007–23.
Duff, Antony, Lindsay Farmer, Sandra Marshall and Victor Tadros (2007) *The Trial on Trial: Volume 3*. Hart, Oxford.
Dutant, Julien (2016) 'How to Be an Infallibilist' *Philosophical Issues* 26(1):148–71.
Easwaran, Kenny, Luke Fenton-Glynn, Christopher Hitchcock, and Joel Velasco (2016) 'Updating on the Credences of Others: Disagreement, Agreement, and Synergy' *Philosophers' Imprint* 16:1–39.

Enoch, David and Talia Fisher (2013) 'Sense And "Sensitivity": Epistemic and Instrumental Approaches to Statistical Evidence' *Stanford Law Review* 67:557–611.

________, Levi Spectre, and Talia Fisher (2012) 'Statistical Evidence, Sensitivity, and the Legal Value of Knowledge' *Philosophy and Public Affairs* 40(3):197–224.

Fantl, Jeremy and Matthew McGrath (2009) *Knowledge in an Uncertain World*. OUP.

Fitelson, Branden and David Jehle (2009) 'What is the "Equal Weight View"?' *Episteme* 6(3):280–293.

Fumerton, Richard (1995) *Metaepistemology and Scepticism* Rowman and Littlefield.

Gardiner, Georgi (2017a) 'In Defence of Reasonable Doubt' *Journal of Applied Philosophy* 34(2):221–241.

________ (2017b) 'Safety's Swamp: Against the Value of Modal Stability' *American Philosophical Quarterly* 54(2):119–129.

________ (2018) 'Legal Burdens of Proof and Statistical Evidence' *The Routledge Handbook of Applied Epistemology* David Coady and James Chase (eds.) Routledge, 171–195.

________ (2019) 'Legal Epistemology' *Oxford Bibliographies: Philosophy* Duncan Pritchard (ed.) OUP.

________ (forthcoming) 'Legal Evidence and Knowledge' *The Routledge Handbook of the Philosophy of Evidence* Maria Lasonen-Aarnio and Clayton Littlejohn (eds.) Routledge.

________ (ms-a) 'Doubt and Disagreement in the #MeToo Era'.

________ (ms-b) 'Profiling and Proof: Are Statistics Safe?'.

________ (ms-c) 'Relevance and Risk: How the Relevant Alternatives Framework Models the Epistemology of Risk'.

________ (ms-d) 'She Said, He Said: Rape Accusations and the Preponderance of Evidence'.

Gerken, Mikkel (2017) *On Folk Epistemology* OUP.

Goldman, Alvin (1976) 'Discrimination and Perceptual Knowledge' *Journal of Philosophy* 73:771–791.

________ (1986) *Epistemology and Cognition* Harvard University Press.

________ and Dennis Whitcomb (2011) *Social Epistemology: Essential Readings* OUP.

Haack, Susan (2014) *Evidence Matters* CUP.

Hawthorne, John (2004) *Knowledge and Lotteries* Clarendon Press.

Ho, Hock Lai (2008) *A Philosophy of Evidence Law* OUP.

________ (2015) 'The Legal Concept of Evidence' *Stanford Encyclopedia of Philosophy* Edward Zalta (ed).

Ichikawa, Jonathan Jenkins (2017) *Contextualising Knowledge* OUP.

Jackson, Elizabeth (2018) 'Belief, Credence, and Evidence' *Synthese*.

Keynes, John Maynard (1921) *A Treatise on Probability* MacMillan.

King, Zoe Johnson (ms) 'The Trouble with Standards of Proof.'

Laudan, Larry (2011) 'The Rules of Trial, Political Morality and the Costs of Error: Or, Is Proof Beyond a Reasonable Doubt Doing More Harm Than Good?' Leslie Green and Brian Leiter (eds.) *Oxford Studies in Philosophy of Law* OUP, 195–227.

Lawlor, Krista (2013) *Assurance* OUP.

Lewis, David (1996) 'Elusive Knowledge' *Australasian Journal of Philosophy* 74:549–567.

List, Christian and Philip Pettit (2011) *Group Agency* OUP.

Littlejohn, Clayton (2018) 'Truth, Knowledge, and the Standard of Proof in Criminal Law' *Synthese*.

McCauliff, C.M.A. (1982) 'Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees? *Vanderbilt Law Review* 35:1293–1335.

McKinnon, Rachel (2013) 'Lotteries, Knowledge, and Irrelevant Alternatives' *Dialogue* 52(3):523–549.

Moss, Sarah (2018a) 'Moral Encroachment' *Proceedings of the Aristotelian Society* 118(2):177–205.

_______ (2018b) *Probabilistic Knowledge* OUP.

Mulrine, Thomas (1997) 'Reasonable Doubt: How in the World Is It Defined?' *American University Journal of International Law and Policy* 12(1):195–225.
Nance, Dale (2016) *The Burdens of Proof* CUP.
Nelkin, Dana (2000) 'The Lottery Paradox, Knowledge, and Rationality' *Philosophical Review* 109(3):373–409.
Pardo, Michael (2019) 'The Paradoxes of Legal Proof: A Critical Guide' *Boston University Law Review*.
Pritchard, Duncan (2005) *Epistemic Luck* OUP.
________ (2018) 'Legal Risk, Legal Evidence and the Arithmetic of Criminal Justice' *Jurisprudence* 9(1):108–119.
Pryor, Jim. (2004) 'What's wrong with Moore's Argument?' *Philosophical Issues* 14:349–378.
Redmayne, Mike (2008) 'Exploring the Proof Paradoxes' *Legal Theory* 14:281–309.
Roberts, Paul (ms) 'Epistemic Standards in English Criminal Jurisprudence.'
Russell, Jeffrey Sanford, John Hawthorne and Lara Buchak (2015) 'Groupthink' *Philosophical Studies* 172(5):1287–1309.
Rysiew, Patrick (2006) 'Motivating the Relevant Alternatives Approach' *Canadian Journal of Philosophy* 36(2):259–279.
Sand, Leonard and Danielle Rose (2003) 'Proof Beyond All Possible Doubt: Is There a Need for a Higher Burden of Proof when the Sentence may be Death?' *Chicago-Kent Law Review* 78:1359–75.
Smith, Martin (2010) 'What Else Justification Could Be?' *Noûs* 44(1):10–31.
________ (2016) *Between Probability and Certainty*: *What Justifies Belief* OUP.
Staffel, Julia (2016) 'Beliefs, Buses and Lotteries: Why Rational Belief Can't Be Stably High Credence' *Philosophical Studies* 173(7):1721–1734.
Starkie, Thomas (1833) *Practical Treatise on the Law of Evidence* Clarke.
Stein, Alex (2005) *Foundations of Evidence Law* OUP.
Stine, Gail (1976) 'Skepticism, Relevant Alternatives, and Deductive Closure' *Philosophical Studies* 29:249–261.
Thomson, Judith Jarvis (1986) 'Liability and Individualized Evidence' *Law and Contemporary Problems* 49(3):199–219.
Tribe, Laurence (1971) 'Trial by Mathematics: Precision and Ritual in the Legal Process' *Harvard Law Review* 84(6):1329–93.