

Knowledge and Legal Proof

Sarah Moss

ssmoss@umich.edu

1 An overview of the knowledge account

What does it take to prove guilt beyond a reasonable doubt? Consider the following example from NESSON 1979, well-known among scholars of evidence:

Prisoners: 25 prisoners are exercising in a prison yard, when 24 of them suddenly join together in a planned attack on the prison guards. The remaining prisoner tries to stop the attack. There is no available evidence distinguishing the innocent prisoner from the rest. Local prosecutors randomly select one of the prisoners and bring him to trial for participating in the attack.

It is widely agreed that the merely statistical evidence in *Prisoners* cannot sustain a verdict of guilt. But why not? After all, the evidence justifies a high degree of confidence in guilt. On the basis of this evidence, it would be reasonable to have .96 credence that the defendant participated in the attack. Increase the numbers, and the same evidence may justify having an arbitrarily high credence that the defendant is guilty. Yet something will be missing—something, apart from mere confidence in guilt, that we must have in order to convict.

Now consider the following example derived from HARMAN 1986, well-known among epistemologists:

Lottery: Jones has purchased a ticket in a small lottery. There are 25 tickets in all. The winning number has already been selected at random, but it hasn't been announced yet.

1. An early precursor of this paper was discussed at the 2017 Analytic Legal Philosophy Conference, University of Texas–Austin, the Rutgers Value Theory Workshop, the Michigan Legal Theory Workshop, the UCLA Legal Theory Workshop, and the book symposia for *Probabilistic Knowledge* held at University of Hamburg and at King's College London. Thanks to these audiences for helpful comments. Thanks also to Nico Cornell, Marcello Di Bello, Mark Greenberg, Scott Hershovitz, Don Herzog, Gabe Mendlow, Michael Pardo, and Eric Swanson for thoughtful feedback on earlier drafts.

Before the winning number is announced, Jones is justified in being confident that she lost the lottery. Increase the size of the lottery, and she may be justified in having an arbitrarily high credence that she lost. But there will still be one respect in which her mental state will fall short, epistemically speaking. As epistemologists are fond of pointing out, as long as Jones can't rule out the possibility that she has the winning ticket, she cannot *know* that she lost the lottery.

Prisoners demonstrates that justified high confidence in a proposition does not suffice for proving it beyond a reasonable doubt. *Lottery* demonstrates that justified high confidence in a proposition does not suffice for knowing it. The examples are strikingly similar. In this paper, I defend the most straightforward explanation for their similarity. At a first pass, my central thesis is just this: *legal proof requires knowledge*. Conviction requires proving beyond a reasonable doubt that the defendant is guilty, and this conclusion is proved if and only if the judge or jury knows it. Since the jury in *Prisoners* can't rule out the possibility that the defendant is the innocent prisoner who tried to stop the attack, they don't know that he is guilty, and that is why it is improper for them to convict.²

Right off the bat, the idea that legal proof requires knowledge has a lot to recommend it. Besides the striking similarity between *Prisoners* and *Lottery*, there is something unsettling about a jury deciding, "For all we know, the defendant might be innocent. But let's vote to convict him anyway." It would sound odd for them to say, "The defendant might be innocent. We find him guilty." When a jury delivers a guilty verdict, they are engaging in a speech act similar to declaring or asserting that the defendant is guilty. As THOMSON 1986 points out, "their agreeing to say those words is not made acceptable by the mere fact that the defendant actually is guilty of what he is charged with" (213). In other words, mere truth is not the correct norm of assertion at a criminal trial, just as it is not the correct norm of assertion, more generally speaking. A more promising proposal is that knowledge is the norm of assertion, including legal assertions of guilt.³

If the knowledge account of legal proof has a lot to recommend it, why isn't it more widely accepted?⁴ There are several stumbling blocks for the knowledge

2. A second pass: strictly speaking, a guilty verdict is proper only if the factfinder knows that the defendant is guilty *on the basis of the evidence presented at trial*. The criminal standard of proof is not met when a judge happens to arrive at the courthouse knowing that the defendant is guilty, if no evidence of guilt is ever given in court. For ease of exposition, I will take this qualification to be implicit throughout.

3. For a canonical defense of the knowledge norm of assertion, see chapter 11 of WILLIAMSON 2000. Additional proponents of the norm include MOORE 1962, UNGER 1975, DEROSE 2002, HAWTHORNE 2004, and STANLEY 2005.

4. There are a few notable exceptions. DUFF et al. 2007 and PARDO 2005 accept that some instances of legal proof require knowledge; in §3 of this paper, I contrast their accounts with mine. Knowledge also plays an interesting role in accounts of legal proof defended by LITTLEJOHN 2017 and BLOME-TILLMANN 2017, though these accounts are much further from mine.

account, objections that may already be on the minds of discerning readers. For starters, it might seem as if proof of guilt does not require jurors to have *knowledge-level* confidence in guilt, but merely to believe that the likelihood of guilt is beyond some high threshold. According to this objection, the knowledge account imposes an overly strict demand on the subjective state of the factfinder—that is, the judge or jury. A second objection is that the knowledge account seems to impose an overly strict *objective* demand on legal proof. Knowledge is factive, i.e., only true contents can be known. If someone is innocent, you can't know that they are guilty; therefore, on the knowledge account, you can't convict them. Yet at first glance, it might seem obvious that our criminal standard of proof permits false convictions, provided they are justified by the evidence. Furthermore, factive standards of proof might appear ill-suited to guide jurors in their decision making. If it is hard for a jury to tell whether a defendant is innocent, then it may also be hard for them to tell whether they *know* that a defendant is innocent. And so, on the knowledge account, it will be hard for them to tell whether the criminal standard of proof is met. But then what good is the criminal standard of proof, if the jury can't tell whether it is satisfied?

This paper answers each of the foregoing objections, as well as several others. In addition, the main part of this paper sets out a positive case for the knowledge account, arguing that it provides illuminating answers to widely debated questions in the theory of evidence. In section 2, I use the knowledge account to explain why the standard of proof beyond a reasonable doubt seems to elude precise definition. In section 3, I discuss why statistical evidence is generally insufficient for legal proof, and why courts are sometimes reluctant even to admit statistical evidence at trial. I present objections to several leading accounts of legal proof, arguing that they make unacceptable predictions about when statistical evidence suffices for proof. I answer some objections to my account along the way, and I address remaining objections in the fourth and final section.

This paper has two limits that I should flag up front. First, my primary aim in this paper is to build a positive case for the knowledge account. I discuss the most prominent alternatives to my account, and I explain why they have problems. But my discussion of existing accounts is not comprehensive. A more thorough defense of my account would include an encyclopedic survey of all existing accounts of legal proof, along with my arguments against each. Second, this paper is more theoretical than prescriptive. I am arguing that a conclusion is legally proved if and only if it is knowledge, but I am not defending substantive claims about what sort of evidence is necessary and sufficient for knowledge. For example, I do not defend any claims about exactly when eyewitness testimony is capable of generating knowledge. This

paper calls attention to structural features of knowledge that make it especially apt for defining what legal proof requires. A recurring theme of the paper is that legal proof requires ruling out relevant epistemic possibilities, where the relevance of a possibility is not determined by its probability. Since this same feature distinguishes knowledge from mere justified confidence, our discussion of it will help to illuminate why legal proof requires the former attitude, rather than the latter.

2 The elusiveness of reasonable doubt

2.1 Attempts to define the reasonable doubt standard

The reasonable doubt standard has long served as the standard of proof in common law criminal trials.⁵ For just as long as it has been around, the standard has been notoriously difficult to define. As Wigmore observes in his *Treatise on Evidence*, “when anything more than a simple caution and a brief definition is given. . . the actual effect upon the jury, instead of being enlightenment, is rather confusion or at least continued incomprehension.”⁶ The United States Supreme Court has held that “[a]ttempts to explain the term ‘reasonable doubt’ do not result in making it any clearer to the minds of the jury.”⁷ The difficulty of explaining the term has not deterred judges from trying, though, and over the past several decades, there has been a gradual accumulation of “a considerable corpus of case law in which judges flounder unhappily over the definition of ‘reasonable doubt,’”⁸ often leading to verdicts that are overturned on appeal. A number of appellate courts are not merely pessimistic about defining reasonable doubt, but openly hostile to the project. In some jurisdictions, giving any definition constitutes grounds for appeal. To quote one representative opinion from the Oklahoma Court of Criminal Appeals, “We are at a loss to understand why trial courts in this jurisdiction continue to give such an instruction when we have condemned them from territorial days to the present. . . it is error for the trial judge to try to define reasonable doubt.”⁹

Given the importance of the criminal standard of proof, it is disconcerting, to say the least, that it often goes without definition. As Judge Posner puts it, “The question whether the prosecution has proved the defendant guilty beyond a reasonable doubt

5. Although the reasonable doubt standard dates from the late eighteenth century, it was most recently clarified by the United States Supreme Court in *In re Winship*, 397 U.S. 358, 364 (1970). According to *Winship*, due process requires that each individual element of a crime be proved beyond a reasonable doubt. For simplicity, I talk about what it takes to prove that a defendant is guilty of a crime, but my arguments extend to the act of proving the individual elements that constitute this claim.

6. WIGMORE 1923, §2497

7. *Holland v. United States*, 348 U.S. 121, 140 (1954) (quoting *Miles v. United States*, 103 U.S. 304, 312 (1880))

8. WHITMAN 2008, p.2

9. *Jones v. Oklahoma*, 554 P.2d 830, 835 (1976)

is central to every criminal trial. Can it be that the term should *never* be defined? Is it a mystical term, a talisman, somehow tarnished by attempts at definition?"¹⁰ This exasperated question is not merely rhetorical, but illuminating. It is not just that it happens to be hard to figure out what it takes to prove something beyond a reasonable doubt. There seems to be something elusive about the standard itself, so that by its very nature, it is nearly impossible to define. As another author puts it,

This debate over definitions has raged as long as the term has been prominent. Although some might argue that the difficulty of the endeavor does not mean that it should be abandoned, "difficulty" is perhaps not the issue. Rather, the issue is the reasonable doubt concept's intrinsic incompatibility with the project of definition. The failure of the legal community to reach a consensus over the last one hundred years suggests that the nature of the concept defies attempts at more precise definition.¹¹

Even the drafters of the Model Penal Code—which is, generally speaking, a wellspring of definitions—decline to define 'reasonable doubt' on the grounds that "definition can add nothing helpful to the phrase."¹²

Hence in addition to the traditional question—*what does it take to prove something beyond a reasonable doubt?*—there is a second question that deserves attention in its own right: *why is this first question so hard to answer?* Why is the criminal standard so elusive? Why have judges and scholars repeatedly concluded that when it comes to defining this standard for jurors, it is almost always better to say less rather than more? There have been few attempts to answer these questions. As Ho 2008 notes, "It is a mystery why a doctrine held in such high esteem should yet be so elusive" (455).

2.2 Explaining the elusiveness of the standard

To solve this mystery, it is helpful at this point to look at an analogous problem faced by the student of the theory of knowledge.¹³ As LEWIS 1996 famously puts it, knowledge is *elusive*. To give a simple example, consider the following attempt to ascribe some ordinary knowledge:

It does sometimes happen that cars get stolen from driveways. But let's ignore that possibility. Now, we can truly say that you know that your car is in your driveway. After all, you can rule out plenty of ordinary possibilities in which your car isn't in your driveway, like the possibility that you parked it on the street.¹⁴

10. *United States v. Hall*, 854 F.2d 1036, 1043 (1988)

11. Note, *Reasonable Doubt: An Argument Against Definition*, 108 Harvard Law Review 1955, 1970 (1995)

12. Brief for the United States at 9, *Victor v. Nebraska*, 114 S.Ct. 1239 (1994), quoting MODEL PENAL CODE cmt. 190 (1985)

13. With apologies to Judith Jarvis Thomson. See THOMSON 1986, p.206.

14. The car theft example is due to VOGEL 1990; it also serves as a focal point of the influential discussion of lottery knowledge in HAWTHORNE 2004.

Having reflected on the possibility that your car has been stolen, it is difficult to say that you know it is in your driveway. FOGELIN 2000 calls this the problem of “epistemic self-destruction,” saying that “some sense should be made of this tendency of epistemic commitments to contract, perhaps to a near-vanishing point, under sustained philosophical scrutiny” (45). According to Lewis, the correct diagnosis of the problem is that philosophical scrutiny raises the epistemic standards for interpreting ‘knowledge’ at our context, where epistemic standards can be understood in terms of *relevant alternatives*—roughly speaking, possibilities that must be ruled out in order for a subject to count as knowing a proposition. By discussing possibilities, we make them relevant. As more possibilities become relevant, our epistemic standards become stricter, and it becomes harder for us to assert true knowledge ascriptions.¹⁵

Out of the epistemology classroom, back to the jury room. The knowledge account of legal proof connects the elusiveness of knowledge with the elusiveness of proof beyond a reasonable doubt, using the former to explain the latter. Suppose that Smith is on trial for murder. The prosecution has just presented an extremely compelling case, providing piles of evidence against Smith. Back in the jury room, a wacky juror could nevertheless manage to raise unreasonable doubts. The juror might insist, “Maybe these massive piles of evidence are part of a big conspiracy! Maybe Smith is actually being set up by the government, as part of a master plan hatched by the D.C. elite!” Or the jury might include an undergraduate philosophy major who insists, “We don’t *know* that Smith is guilty! Maybe Smith is innocent, and we are all just victims of an evil demon, who is controlling our minds and feeding us false memories of evidence in order to get us to convict!” The reasonable doubt standard is meant to protect jury deliberation from exactly this sort of epistemic inflation. The standard places an important constraint on the context in which the jury should be deliberating—namely, it must be a context in which they are considering all and only doubts that it is reasonable for them to consider. It is important that jurors be able to act in good conscience when they vote to convict a defendant. The jury should not vote to convict Smith if they actually believe that he might be innocent. At the same time, there will always be far-fetched possibilities that jurors ought to set aside, such as possibilities involving evil mind-controlling demons. As the United States Supreme Court put it in *Hopt v. Utah*, “Persons of speculative minds may in almost every such case suggest possibilities of truth being different from that established by the most convincing proof. The jurors are not to be led away by such speculative

15. I have described one contextualist approach in the text, but the argument of this section does not rely on relevant alternatives contextualism, or indeed on any particular theory of the elusiveness of knowledge. For the purposes of my argument, what matters is simply that knowledge is elusive; given this fact, my knowledge account of proof entails that proof beyond a reasonable doubt is elusive as well.

notions as to such possibilities.”¹⁶ It is crucial that jury deliberations be guided by a reasonable epistemic standard, if there are to be any proper convictions at all.

Now we can see why it is hard to define this reasonable epistemic standard in greater detail. The problem is that the more we say in an effort to spell out the difference between reasonable and unreasonable doubts, the more we call attention to the possibilities that jurors *shouldn't* be considering. As Lewis might have put it, the more we risk “destroying” the knowledge that would have sustained a conviction. For example, consider the following jury instruction:

It does sometimes happen that a defendant is framed as part of a government conspiracy. But you should ignore this possibility, along with any other possibility it would be unreasonable for you to consider. Now, if you can truly say that you know that the defendant is guilty, you should return a guilty verdict.

A simple characterization of the reasonable doubt standard would instruct jurors to deliberate without considering unreasonable possibilities. But by mentioning those possibilities, we risk taking jurors out of the very context that we are telling them to be in.¹⁷ The less that one says about unreasonable doubts, the easier it is for those doubts to be ignored. Insofar as attempts to clarify the reasonable doubt standard may undermine that standard, there is some reason for judges to refrain from clarifying the standard in jury instructions, and there is reason to expect that academic attempts to clarify the standard may be ill-fated.

We have explained why we can expect it to be difficult to define proof beyond a reasonable doubt. Is there anything that one *can* say to explicate this standard? A small number of strategies have been identified by courts as relatively promising. The good news is that most jurors are neither conspiracy theorists nor philosophy majors. As jurors go about their ordinary lives, they routinely set aside far-fetched possibilities and come to know things. As LEWIS 1996 reminds us, “Maybe we do know a lot in daily life; but maybe when we look hard at our knowledge, it goes away. But only when we look at it harder than the sane ever do in daily life” (550). Accordingly, one might hope to clarify the reasonable doubt standard by comparing it with a more familiar epistemic standard. And indeed, in this spirit, courts have explicated the reasonable doubt standard with the following instruction:

Proof beyond a reasonable doubt must, therefore, be proof of such a convincing character that a reasonable person would not hesitate to rely and act upon it in the most important of his own affairs.¹⁸

16. *Hopt v. Utah*, 120 U.S. 430, 440 (1886)

17. If we are not careful, we do the same thing to ourselves in the context of academic discussions of reasonable doubt.

18. *United States v. Savulj*, 700 F.2d 51, 69 (2nd Cir. 1983)

A reasonable person often relies on knowledge in making high-stakes decisions outside the courtroom. According to the above jury instruction, a similar state of knowledge should govern high-stakes decisions inside the courtroom.¹⁹ Alternatively, some courts have managed to clarify the reasonable doubt standard by contrasting it with a stricter one—in particular, with a standard that is so strict that it is practically useless:

Anything can be possible. . . [A] planet could be made out of blue cheese. But that's really not in the realm of what we're talking about.²⁰

It is not a flimsy, fanciful, fictitious doubt which you could raise about anything and everything.²¹

These definitions work well because speakers generally presuppose that their context-sensitive expressions will be interpreted relative to *useful* standards. Although epistemic standards are often subject to inflation, they are somewhat more resilient against the sort of radical inflation that would leave speakers without any useful knowledge ascriptions at all. To sum up, the knowledge account not only explains why most attempts to describe reasonable doubt fail, but also why certain attempts are more likely to succeed.

2.3 Attractive features of my account

There is something surprising about my account of reasonable doubt. From a contemporary perspective, it is natural to assume that the requirement of proof beyond a reasonable doubt is meant to protect the accused against hasty conviction. But on my account, the reasonable doubt standard is intended to guard against epistemic inflation, thereby making it *easier* for jurors to convict. Interestingly enough, my account of the legal function of reasonable doubt reflects the historical origins of this notion. The reasonable doubt standard emerged in the late eighteenth century in response to an unreasonable inflation of epistemic standards in jury trials around that time. Just as our conspiracy theorist is in the grips of ideological message boards, and our philosophy major is in the grips of Descartes' First Meditation, so the jurors of the

19. This discussion raises an important question that I do not have space to address here—namely, can the criminal standard of proof vary depending on what is at stake in a criminal trial? LILLQUIST 2002 argues that “the reasonable doubt standard of proof is inevitably flexible in nature: in some cases juries will require more proof than in other cases” (85); SAND & ROSE 2003 argue that a higher standard should apply in death-eligible cases; LAUDAN & SAUNDERS 2009 argue that “differences in the seriousness of crimes and in the severity of their associated sentences make it questionable whether we should continue to seek a single standard for all seasons instead of a set of standards of varying degrees of severity” (1). These theories are naturally motivated and explained by my account of legal proof.

20. *Victor v. Nebraska*, 511 U.S. 1, 17 (1994), quoted approvingly by Justice O'Connor as an accurate elaboration of the notion of reasonable doubt.

21. *People v. Davis*, 171 Mich. 241, 137 N.W. 61 (1912)

eighteenth century were in the grips of a popular literature of conscience that made them reluctant to deliver any guilty verdicts. As WHITMAN 2008 argues, jurors of the time believed that it was a mortal sin for them to convict an innocent person.²² Accordingly, the potential consequences of false conviction were severe, including “vengeance of God upon his family and trade, body and soul, in this world and that to come.”²³ With their own salvation at stake, jurors were understandably motivated to consider any conceivable possibility that the defendant might be innocent of the crime charged. In 1785, the moral philosopher William Paley describes the result of these heightened epistemic standards:

I apprehend much harm to have been done to the community, by the overstrained scrupulousness, or weak timidity of juries, which demands often such proof of a prisoner’s guilt, as the nature and secrecy of his crime scarce possibly admit of; and which holds it the part of a *safe* conscience not to condemn any man, whilst there exists the minutest possibility of his innocence.²⁴

As Whitman argues, the reasonable doubt standard emerged in this same time period in order to counteract the epistemic inflation described by Paley. The language of reasonable doubt was used to encourage jurors to deliberate using a more reasonable and more useful standard. As one judge put it,

[For almost any crime,] there is a strict possibility, that somebody else might have committed it: But that the nature of evidence requires, that Juries should not govern themselves, in questions of evidence, that come before them, by that strictness, is most evident, for if it were not so, it is not possible that offenders of any kind should be brought to Justice.²⁵

Insofar as the historical origins of the notion of reasonable doubt might shed light on its contemporary legal function, my account of the latter has the advantage of reflecting the former.

A second advantage to my discussion of reasonable doubt is that it can help answer an objection to the knowledge account of legal proof—namely, that the account imposes an overly strict demand on the subjective state of the factfinder. At first glance, one might be tempted to make the following objection:

You do not have to know that the defendant is guilty before you can convict him. It is only necessary that you should believe from the evidence beyond a reasonable doubt, that he is guilty, and if you do so believe from all of the evidence, beyond a reasonable doubt, that the defendant is guilty, then it is your sworn duty to so find.

22. For instance, in a case where a defendant is falsely convicted and condemned to death, it was believed that “the Jury *in foro conscientiae* are certainly guilty of his murder” (HAWLES 1680, p.22).

23. H.E. 1664, p.1

24. PALEY 1785, p.550. For discussion of this passage and its historical context, see GREEN 1985.

25. Trial of Thomas Hornsby (Theft with Violence, Highway Robbery) (1782). The Proceedings of the Old Bailey Ref: T17830430-67

This passage is taken from a jury instruction given by a trial court judge in Mississippi and subsequently condemned by the Mississippi Supreme Court.²⁶ According to this instruction, knowledge requires eliminating *all doubts*, whereas conviction merely requires eliminating *all reasonable doubts*—and hence, conviction cannot require knowledge. There is something compelling about this line of thought, as it is indeed the case that eliminating all doubts means eliminating unreasonable doubts as well as reasonable ones. However, as long as jurors are being reasonable, there will not be any difference from their perspective between eliminating all doubts and eliminating reasonable ones. The objection itself is problematic because by considering it, one calls attention to this difference, and implicitly grants the existence of the very possibilities that the reasonable doubt standard calls jurors to set aside. That is what is wrong with issuing the above jury instruction, and why it was right for the Mississippi Supreme Court to reject it. The instruction takes the jurors out of the very context that they are being instructed to deliberate in.

There is a difference between being reasonable and merely acting as if you are being reasonable. We hold jurors to the former standard. Consider a jury made up entirely of unreasonable conspiracy theorists. No amount of evidence would convince them that the defendant is guilty. Should we encourage these jurors to say, “The defendant might be innocent, but we should convict him anyway”? This might be a decent response to a bad situation, but it wouldn’t be the best response. To make the best of this bad situation does not mean telling the jurors to convict against their better judgment, but rather demanding that they judge reasonably and then render their verdict accordingly. In a similar spirit, consider a jury of ignorant racists who are disposed to tune out any evidence that suggests that a white person is guilty. Should we encourage these jurors to acquit white defendants? In this case, the best response is obvious: jurors are not merely required to return the verdict that is proper given what they believe, but also required to be reasonable as they form their beliefs. In order to be reasonable, jurors must form their beliefs in accordance with the evidence. The reasonable doubt standard imposes a constraint with a similar form: in order to be reasonable, jurors must consider only doubts that it is reasonable for them to consider.²⁷

A third advantage of my account of the reasonable doubt standard is that it sheds

26. *Nobles v. State of Mississippi*, 241 So.2d 826 (1970)

27. This discussion calls attention to one sort of misconduct that the defense might engage in—namely, the act of raising possibilities *in order to prevent jurors from being able to engage in reasonable deliberation*. The correct response to such misconduct is not to instruct jurors to convict someone that they sincerely believe might be innocent. Rather, the correct response to defense misconduct will always depend on a range of pragmatic factors, including the costs of declaring a mistrial and the feasibility of proceeding as if the misconduct had not occurred—as Lewis puts it, whether jurors can “bend the rules of cooperative conversation” so that “they may rightly be said to know that the accused is guilty” (560, 556).

light on fundamental problems with a certain sort of attempt to define the standard. I have argued that the standard is elusive in the sense that attempting to define it often makes it harder to follow. The reasonable doubt standard is also elusive in another very specific sense—namely, the standard is *impossible to quantify*. It is tempting to try to define the standard by throwing out numbers.²⁸ In fact, plenty of lawyers have done just that, making statements such as the following:

I like to make it kind of like a football field where you start at one end and you go to the other. If you go all the way and make a touchdown, that's like a hundred percent. That's beyond no doubt. I like to say reasonable doubt is kind of like 75 percent. Somewhere—75 and 90.²⁹

It's beyond a reasonable doubt, so it's beyond the 50 yard line. You have to take it to the opponent's 20, the red zone. You got to get it in the red zone for beyond a reasonable doubt.³⁰

Appellate courts routinely condemn analogies like these, stating that the criminal standard of proof cannot be defined in terms of any threshold notion of confidence.³¹ As we have seen, the example of *Prisoners* is intended to make just this point. *Prisoners* demonstrates that justified high credence in a proposition does not suffice for proving it beyond a reasonable doubt.

The knowledge account of the criminal standard helps us better understand this fact. It is impossible to quantify reasonable doubt because whether a doubt is *probable* and whether it is *capable of undermining knowledge* are distinct questions. As Posner observes, “It is one thing to tell jurors to set aside unreasonable doubts, another to tell them to determine whether the probability that the defendant is guilty is more than 75, or 95, or 99 percent.”³² Following LEWIS 1996, one might say that knowledge of guilt requires ruling out relevant possibilities in which the defendant is innocent, where the relevance of a possibility may depend on its non-probabilistic structural features. To return to our earlier example, suppose 25 separate trials are conducted in *Prisoners*, one for each prisoner in the yard. If each jury could reasonably set aside the possibility that the defendant is innocent, then our standard of proof would demand 25 convictions and ensure the conviction of an innocent person. Even if false

28. For instance, McCauliff 1982 reports that among 171 federal judges and United States Supreme Court justices, the average probability threshold associated with proof beyond a reasonable doubt was .90. According to Magnusson et al. 2014, trial judges report intentions to return guilty verdicts in mock trial settings when their credences in guilt exceed .83.

29. *State v. Casey*, No. 19940, 2004 WL 405738, at para. 43 (Ohio App. Mar. 5, 2004)

30. *People v. Lewis*, 2015 IL App (1st) 130171, 33 N.E.3d 212, 392 Ill. Dec. 663

31. *Commonwealth v. Rosa*, 661 N.E.2d 56, 63 (Mass. 1996) (“it is best for judges to avoid examples that have numeric or quantifiable implications”); *State v. Rizzo*, 833 A.2d 363, 399 (Conn. 2003) (“it is improper for a trial court to attempt to explain the concept of reasonable doubt by metaphors or analogies that are quantified in nature”).

32. *United States v. Hall*, 854 F.2d 1036, 1044 (7th Cir. 1988) (Posner, J., concurring)

convictions are prevalent in practice, it is counterintuitive to think that they could ever be required in certain circumstances. As ZUCKERMAN 1986 observes, convicting all 25 prisoners would have us “openly acknowledge that the individual defendant may well belong to the innocent minority, and therefore undermine the citizen’s confidence that the legal system will protect him from mistaken conviction of crime or mistaken imposition of liability.” (499). As Lewis would put it, the Rule of Actuality and Rule of Resemblance together constrain the possibilities that jurors may properly ignore. Since one of the prisoners is innocent, and his innocence is just like that of any other prisoner, we must acquit all 25. Hence non-probabilistic features of possibilities can make a difference to their relevance—and from this fact, it follows that the reasonable doubt standard cannot be quantified.

3 Statistical evidence

3.1 The general problem of statistical evidence

So far, we have been focusing on the criminal standard of proof. At this point, we should broaden our discussion to include other standards, such as proof by clear and convincing evidence, proof by a preponderance of the evidence, and so on.³³ In other words, we should turn to discussing the notion of *legal proof* in general.

Although standards of proof differ in strength, they all require the factfinder to have at least some amount of confidence in the conclusion to be proved. As Justice Harlan states, “Although the phrases ‘preponderance of the evidence’ and ‘proof beyond a reasonable doubt’ are quantitatively imprecise, they do communicate to the finder of fact different notions concerning the degree of confidence he is expected to have in the correctness of his factual conclusions.”³⁴ BENTHAM 1843 was among the first to clarify this notion of confidence, stating that legal proof “admits of, and exists in, different degrees of strength,” where “the practice of wagering affords at the same time a proof of the existence, and a mode of expression or measurement for [these] quantities or degrees” (223). In the parlance of contemporary epistemologists, Bentham is talking about credences. The odds at which you are willing to bet on a proposition reflects how much credence you have in it. Proving a fact by a certain standard requires the factfinder to have at least a certain amount of credence in that fact. For instance, proof of liability by a preponderance of the evidence requires the factfinder to have greater than .5 credence that the defendant is liable.

33. See McCAULIFF 1982 for an overview of standards of proof, including several that are less demanding than the preponderance standard. For instance, the standard of reasonable suspicion—applied in stop-and-frisk searches—is commonly associated with a probability threshold around .3.

34. *In re Winship*, 397 U.S. 358, 370 (1970)

This subjective requirement on credences is necessary but not sufficient for legal proof. As we have seen, merely statistical evidence can justify having a certain credence in a claim, without proving it beyond a reasonable doubt. This is another feature common to all standards of proof, as statistical evidence is also insufficient to establish proof by more lenient standards. Consider the following example from COHEN 1977:

Gatecrasher: [I]t is common ground that 499 people paid for admission to a rodeo, and that 1,000 are counted on the seats, of whom *A* is one. Suppose no tickets were issued and there can be no testimony as to whether *A* paid for admission or climbed over the fence. So by any plausible criterion of mathematical probability there is a .501 probability, on the admitted facts, that he did not pay. (74)

Just as in *Prisoners*, the merely statistical evidence in *Gatecrasher* is insufficient to sustain a verdict. As COHEN 1981 puts it, “our intuitions of justice revolt against the idea that the plaintiff should be awarded judgment” (627). The insufficiency of statistical evidence must be explained not by some feature of the reasonable doubt standard in particular, but rather by some feature that all standards of proof have in common. That is, the problem of statistical evidence must be solved by a general account of what legal proof requires. I have argued that the evidence in *Prisoners* cannot prove guilt beyond a reasonable doubt because it cannot provide the factfinder with knowledge. In what follows, I extend the knowledge account to other standards of proof. Generally speaking, statistical evidence fails to provide legal proof because it fails to provide the factfinder with a certain sort of knowledge.

To clarify what I am aiming to explain: our hostility to statistical evidence is not without qualification. Courts have also been increasingly willing to issue convictions on the basis of random match DNA evidence.³⁵ Some courts have been willing to assign liability in civil cases that look remarkably like *Gatecrasher*.³⁶ In addition, courts have been willing to base verdicts on statistical evidence in market-share liability cases. Suppose that a plaintiff proves that multiple manufacturers were selling defective products and that one of these products harmed her, but she cannot prove that any particular manufacturer specifically caused her harm. In *Sindell v. Abbott Laboratories*, the California Supreme Court held that in such cases, the court may assign liability to manufacturers in proportion with their share of the market for the defective product.³⁷ For any company with a small market share, the statistical evidence suggests that it is very unlikely that they are liable for the harm in question. Hence the evidence is obviously insufficient to prove by a preponderance of the evidence

35. See ROTH 2010 for discussion of the recent rise in “pure cold hit” prosecutions.

36. The canonical example is *Kaminsky v. Hertz Corp.* 288 N.W.2d 426 (1979). See §3.5 for further discussion.

37. *Sindell v. Abbott Laboratories* 26 Cal. 3d 588 (1980)

that the small company is liable, as the evidence does not even justify having a high credence that they caused the harm.³⁸ The problem of statistical evidence is not the problem of explaining why statistical evidence is never sufficient for a verdict, but rather why it is insufficient in many normal cases.

As a final note of clarification, I should mention that a small number of legal scholars reject the problem of statistical evidence altogether.³⁹ According to a more revisionary approach to the problem, there is nothing objectionable *per se* about basing verdicts on merely statistical evidence. Courts should aim to maximize the accuracy of verdicts, and any intuitive hostility to highly probative statistical evidence is misguided. For present purposes, I will set aside this approach and focus on the project of making sense of a widely held hostility to statistical evidence, rather than defending that hostility from more basic principles.

3.2 Objections to moral accounts

Before introducing my preferred solution to the problem of statistical evidence, I want to raise some objections to several existing proposals for solving it. Broadly speaking, existing accounts of statistical evidence fall into two categories. According to epistemic accounts, statistical evidence is incapable of providing the sort of proof that non-statistical evidence can provide. According to moral accounts, statistical evidence might turn out to prove conclusions just as well as any other evidence, but it is morally objectionable to act on those conclusions, or to use statistical evidence to prove them.

According to the moral account defended by ZUCKERMAN 1986, punishing a defendant on the basis of statistical evidence is tantamount to punishing him for being a member of a group, most of whose members are guilty of a crime. Our moral and legal values resist verdicts based on statistical evidence in just the same way that they resist acts of corporate punishment, acts that hold an entire social group responsible for the transgressions of its individual members (499). In a similar spirit, COHEN 1987 argues, “A person who deliberately runs his life in such a way as not to commit torts or break contracts is not to be put at risk by the probative procedures of the system just because he falls into a category of which the majority happen to be tort-feasors or contract-breakers” (94). Both Zuckerman and Cohen are worried about the moral permissibility of acting on conclusions supported by statistical inferences. Another

38. Defendants have also been assigned market-share liability after introducing decisive evidence that they did not cause a specific harm; see *Hymowitz v. Eli Lilly & Co.*, 73 N.Y.2d 487, 541 N.Y.S.2d 941, 539 N.E.2d 1069 (1989). This further supports the conclusion that the doctrine of market-share liability constitutes a *sui generis* exception to the typical burden of proving causation by a preponderance of the evidence. For additional arguments in support of this conclusion, see STEEL 2015.

39. For sympathetic discussion of this approach, see SAKS & KIDD 1980, SHAVIRO 1989, and SCHAUER 2003.

moral account is defended by WASSERMAN 1991, who argues that even just the practice of using statistical evidence to form beliefs about an individual is itself morally problematic:

[W]hat is objectionable is the reliance on others' conduct, or the defendant's past conduct, to infer his commission of a wrongful act. We object to this inference because it ignores the defendant's capacity to diverge from his associates or from his past, thereby demeaning his individuality... (942-3)

All of these accounts raise moral concerns related to inferences like the following:

- (1) a. Most *F*s are tort-feasors.
- b. *X* is an *F*.
- c. Therefore, it is more likely than not that *X* is a tort-feasor.

As these authors see it, an inference of this form may be perfectly capable of proving its conclusion. But the inference itself is morally suspect, or it is morally incapable of grounding a verdict. According to Wasserman, someone who infers as in (1) is failing to treat *X* as an individual. According to Zuckerman and Cohen, basing a verdict on (1) is tantamount to assigning liability to *X* *merely because X is an F*, as opposed to assigning liability for some morally legitimate reason, such as the fact that *X* probably caused some particular harm.

These moral accounts are tempting. As we will later see, there is something right about them. But unfortunately, they do not solve the problem of statistical evidence, because they are not broad enough to explain our general hostility to verdicts based on statistical inferences. The accounts explain our judgments about *Prisoners* and *Gatecrasher*, but only because these examples have certain idiosyncratic features—namely, both involve profiling an individual person on the basis of his membership in a group that consists mainly of bad actors. But many other sorts of statistical inferences are just as incapable of providing legal proof. The subject of a problematic statistical inference need not be the defendant. The predicate need not be any negative property. Consider the following example:

Dog Bite: There are ten dogs that roam the neighborhood—six owned by Jones, and four owned by Smith. The local pet ordinances do not require owners to keep their dogs on a leash, but owners are liable for any harm caused by dogs that roam freely. Little Bobby sees ten dogs in a local park. He randomly picks one to play with, and the dog bites him. It can be established that the dogs in the park were all and only the neighborhood dogs. But Bobby is too young to tell the dogs apart, so there is no available evidence as to which dog bit him.

Given the available evidence, it is reasonable to have .6 credence that the dog that bit Bobby was owned by Jones. But it is impermissible to hold Jones responsible for the dog bite, merely on the grounds that she happens to own a majority of the dogs in

the neighborhood. Just as in *Gatecrasher*, the merely statistical evidence in *Dog Bite* is insufficient to sustain a verdict. Yet the statistical inference in *Dog Bite* has the following form:

- (2) a. Most of the neighborhood dogs are owned by Jones.
- b. The dog that bit Bobby is a neighborhood dog.
- c. Therefore, it is more likely than not that the dog that bit Bobby is owned by Jones.

Whose individuality is demeaned by this inference? The subject of the inference is a dog. Jones is not being treated like an arbitrary dog owner; indeed, since there are only two dog owners in the neighborhood, it would be impossible to unfairly assign her liability on the basis of the fact that most local dog owners have some property.

Suppose we ask Jones why we shouldn't hold her responsible for the dog bite. She would be right to respond, "Look, the mere fact that I own most of the dogs in the neighborhood doesn't *prove anything*." And in a sense, she's right. The problem in *Dog Bite* is not that we have made some inference we shouldn't have. The problem is that there is an epistemic shortcoming to the case against Jones. Let us turn to epistemic accounts of statistical evidence, then, to see if they can do any better.

3.3 Objections to epistemic accounts

According to most leading epistemic accounts of statistical evidence, something more than justification is required by legal proof. There are a number of epistemic features that could be the missing ingredient. THOMSON 1986 defends a causal account, arguing that legal proof requires evidence that is "in an appropriate way causally connected with the fact that the defendant caused the harm" (203), where two facts are causally connected just in case one causes the other or they both share a common cause. For example, the testimony of an eyewitness can prove that a defendant is liable for trespass, since the fact that the defendant trespassed is a cause of her testimony. By contrast, the fact that most rodeo-goers trespassed is causally unconnected with the fact that any particular defendant trespassed, and so the former fact is insufficient to prove the latter, even by a preponderance of the evidence.⁴⁰

According to a second epistemic account defended in ENOCH et al. 2012, a verdict must be *sensitive* in the following sense:

For A's belief that p to be sensitive is for it to be the case that had p been false,

⁴⁰ Another causal account is defended by SORENSEN 2006, who argues that a crime justifies a verdict only by being a cause of that verdict—namely, by causing some evidence that causes the verdict (170).

A would probably not have believed that p . (210)⁴¹

Enoch et al. do not give a precise statement of their account. But they do say that they try to give an explanation of “the reluctance to rely on statistical evidence. . . in terms of Sensitivity” (220). They explain their account by contrasting statistical evidence with reliable testimony:

Our eyewitness is not infallible, of course, but she is pretty reliable, and so had it not been a Blue Bus bus, she would have probably not testified that it was; and in that case we would not have found the Blue Bus Company liable. So in this scenario, the finding is appropriately sensitive. Things are different, though, if we base our finding solely on statistical evidence. (206)

Hence it is natural to interpret Enoch et al. as endorsing the following thesis:

(Sensitive Verdict) A finding of liability is proper only if the following condition holds: if the defendant hadn’t been liable, the factfinder probably wouldn’t have found him liable.

Enoch et al. implicate that they accept the analogous claim for criminal verdicts: proper conviction requires that if the defendant had been innocent, the factfinder probably wouldn’t have found him guilty.⁴²

Before I raise my main objections to these epistemic accounts, a quick detour is necessary. There is one technical problem with the above sensitivity account. A lot of verdicts based on statistical evidence actually happen to be sensitive, but for trivial reasons. For instance, consider *Dog Bite*. Let us imagine that as a matter of fact, Bobby was bitten by a dog owned by Jones. Suppose that Jones hadn’t been liable for any harm to Bobby. What would probably have been the case? Well, in the vast majority of scenarios in which Jones isn’t liable for any harm to Bobby, that’s because Bobby *isn’t harmed at all*. After all, it’s not as if the neighborhood dogs were all lining up to bite Bobby, and if he hadn’t been bit by a dog owned by Jones, he would have been bit by a dog owned by Smith. Suppose the jurors find Jones liable for causing harm to Bobby. Their verdict is going to be sensitive to the fact that she is liable, since if Jones hadn’t been liable, the jurors probably wouldn’t have found her liable—namely, because there probably wouldn’t have been a trial at all.

We need to fix up the sensitivity account to avoid this bad result. On behalf of Enoch et al., I propose the following: say that an evidence proposition p is sensitive

41. This condition is inspired by the definition of sensitive belief in NOZICK 1981. As applied to verdicts, the condition says that the verdict would probably not have been issued if its content had been false.

42. Enoch et al. alternate between saying that the jury must *deliver sensitive verdicts* and that they must *have sensitive beliefs*. The latter vocabulary suggests a slight variation on the account given above—namely, that a finding of liability requires that if the defendant hadn’t been liable, the factfinder probably wouldn’t have believed that he was liable. This variation does not make any difference to my arguments.

to the defendant's liability if and only if it satisfies the following condition: if the defendant hadn't been liable, then p would probably have been false. Rather than saying that legal verdicts must be sensitive, let us say that legal verdicts must be based on sensitive evidence:

(Sensitive Evidence) A finding of liability is proper only if the following condition holds: the factfinder believes that the defendant is probably liable, and they base their belief on at least one evidence proposition that is sensitive to his liability.

This version of the sensitivity account delivers the right verdict about *Dog Bite*. The jurors' belief that Jones is probably liable is based entirely on the evidence proposition that she owns most of the dogs in the neighborhood, and this evidence proposition is not sensitive to the fact that she is liable. By contrast, consider a case where Jones is found liable on the basis of reliable eyewitness testimony. In this case, the jurors' belief that Jones is probably liable is based on the content of the testimony offered by the eyewitness—namely, the proposition *that she saw one of Jones's dogs bite Bobby*. This evidence proposition does counterfactually depend on the dog biting Bobby, so the relevant sensitivity condition is satisfied.

Unfortunately, serious problems remain. Several examples make trouble for both the causation account and the sensitivity account of statistical evidence. In a number of cases, statistical evidence is insufficient to sustain a verdict, even though it is causally connected with the fact to be proved and also sensitive to that fact. For instance, sometimes the statistical evidence causes the fact to be proved:

Reluctant Prisoner: 50 prisoners are in a prison yard. Some of them start a riot, and others start to join in. An especially reluctant prisoner decides that he will only join the riot if at least 48 others participate. Eventually this happens, and he joins in, leaving just one innocent prisoner who refuses to join the riot. Local prosecutors randomly select one of the prisoners from the yard and bring him to trial. By sheer coincidence, the randomly selected prisoner is the reluctant prisoner, the last to join the riot.

In other cases, the fact to be proved causes the statistical evidence:

Bold Prisoner: 25 prisoners are in a prison yard. An especially bold prisoner attacks the prison guards and thereby causes a riot. 23 prisoners join in the riot, while the remaining prisoner tries to stop it. Local prosecutors randomly select one of the prisoners from the yard and bring him to trial. By sheer coincidence, the randomly selected prisoner is the bold prisoner responsible for the riot.⁴³

Finally, the statistical evidence and the fact to be proved can have a common cause:

43. For a structurally similar example, see §4 of BLOME-TILLMANN 2015.

Playing Cards: The local prison has four teams of prison guards and thirteen buildings, with exactly 25 prisoners living in each building. Every afternoon, a playing card is drawn at random. The suit of the card determines which guard team is on yard duty, and the number determines which building of prisoners goes out.

All the prisoners hate the Team Diamond guards. They decide that there should be a riot in the prison yard on the next day that Team Diamond has yard duty. That afternoon, the five of diamonds is drawn. Team Diamond and Building Five enter the yard. 24 of the prisoners riot, while the remaining prisoner tries to stop them. Local prosecutors randomly select one of the Building Five prisoners and bring him to trial. As it happens, the selected prisoner was indeed part of the riot.

In each of these examples, the jury has statistical evidence that the defendant is guilty—namely, they know that the defendant was in the yard and that at least 96 percent of the prisoners in the yard are guilty. Furthermore, this evidence is not misleading; the defendant is indeed guilty of attacking the guards. In each case, this merely statistical evidence is insufficient to sustain a verdict of guilt. Unfortunately, the causal account of statistical evidence fails to deliver this result. As noted above, all of these examples involve causal connections obtaining between the statistical evidence and the fact to be proved.

Turning to sensitivity accounts, we can ask: if the defendant hadn't attacked the guards, what would probably have been the case? In two of the three cases, we get the wrong answer. In *Bold Prisoner*, the defendant caused the riot. We can suppose that if he hadn't started the riot, the guards wouldn't have been harmed at all. Hence the jurors are basing their verdict on evidence that is sensitive to the fact that the defendant is guilty, in virtue of being causally connected to it. In the case of *Playing Cards*, any of the other twelve buildings could easily have been selected to enter the prison yard. In the vast majority of counterfactual scenarios where the defendant isn't guilty, that's because some other card was picked, and some other prisoners entered the yard and rioted against the Team Diamond guards. In any such scenario, the evidence proposition that the actual defendant entered the yard would have been false. Hence the jurors in *Playing Cards* also base their guilty verdict on an evidence proposition that is sensitive to the guilt of the defendant. To sum up, the fact that statistical evidence is incapable of sustaining a verdict cannot be grounded in its lack of causal connection or in its insensitivity to the facts.

As I have interpreted the causal and sensitivity accounts, they each aim to state a necessary condition on evidence such that: (a) our total evidence must satisfy the condition in order for it to be sufficient for a verdict, and (b) any total evidence consisting merely of statistical facts fails to satisfy the condition. The foregoing examples challenge the second half of each account, demonstrating that some statistical evidence actually satisfies the proposed necessary conditions of causal connection and sensi-

tivity. Some additional examples challenge the former half of each account, demonstrating that evidence can be sufficient for a verdict, even without being sensitive to the verdict or bearing any substantial causal connection to it. Here is an example of sufficiency without sensitivity:

Grand Canyon: Acme Corp. is charged with dumping waste onto federal land without a permit. At trial, the prosecution presents video evidence of Acme employees heaving a trash bag into a chute that dumps out into the Grand Canyon.⁴⁴

The video evidence in *Pollution* is intuitively sufficient for a verdict. But it is easy enough to fill out the details of the case such that if Acme Corp. had not in fact been guilty of dumping waste onto federal land, it would have been because the trash bag in the video had gotten stuck in the chute and remained on their private property. But in that case, the jury would have been presented with the very same evidence as they actually have. The video evidence is sufficient for legal proof without being sensitive to the fact that it proves.

Evidence can also be sufficient for a guilty verdict without bearing any substantial causal connection to the guilt of the defendant. Consider the following example:

Spit in the Sink: Alice is found murdered in her bedroom. A detailed forensic study proves the following facts: (a) the murder took place in her apartment within the last month, (b) no one except for Alice entered her apartment during that time, except for one person who left some spit her her bathroom sink, and (c) that spit came from the defendant.⁴⁵

Together with the fact that no one else entered her apartment, the presence of the spit in the sink is evidence that the defendant killed Alice. Assuming the forensic science is trustworthy enough, this evidence could sustain a guilty verdict. But what causal connection holds between the fact that the defendant killed Alice, and the fact that he left spit in her sink? Perhaps the murder was caused by greed and jealousy, while the spit was caused by a sneeze, which was itself caused by a cold virus. The murder and the spit have a common cause in only a very attenuated sense—namely, the event of the defendant entering Alice’s apartment is a causal enabling condition of each. But in this attenuated sense, many pairs of events count as having a common cause. In *Prisoners*, the event of a certain group of 25 prisoners entering the yard is an enabling condition of: (a) the event of 24 of those prisoners attacking the guard, and also (b)

44. This case is modeled on the trash chute case in Sosa 1999, offered by Sosa as a counterexample to a sensitivity requirement on knowledge.

45. For simplicity, this example involves some science fiction. Actual forensic science is not good enough to prove these facts. To flesh out the example, suppose that each person necessarily has a radically different gut biome, and that this has enabled the construction of a National Spit Registry that can link any sample of spit with its unique source.

the event of the defendant attacking the guard. Yet this shared enabling condition does not render (a) capable of proving (b). This sort of attenuated causal connection is not useful for saying what it takes for evidence to be sufficient for a verdict.

3.4 Explaining the insufficiency of statistical evidence

I have described several problems for existing accounts of statistical evidence. Happily, where other accounts fail, the knowledge account succeeds. The knowledge account does not rely on any assumptions about the form of statistical inferences that fail to sustain verdicts. The statistical evidence in *Dog Bite* may be just as incapable of grounding knowledge as the statistical evidence in *Gatecrasher*. The examples *Bold Prisoner*, *Reluctant Prisoner*, and *Playing Cards* are all paradigmatic lottery cases, featuring evidence that fails to provide the jury with knowledge that the defendant is guilty. The knowledge account correctly predicts that the evidence in these cases is insufficient to sustain a guilty verdict. By contrast, the factfinder in *Grand Canyon* can know on the basis of video evidence that Acme dumped waste onto federal land, and the factfinder in *Spit in the Sink* can know on the basis of the forensic evidence that the defendant murdered Alice. Hence the knowledge account correctly predicts that the evidence in these cases is sufficient to sustain a guilty verdict.

In addition to solving problems faced by existing accounts, the knowledge account explains a recurring observation in the literature on statistical evidence—namely, that legal proof seems to require *something that looks an awful lot like knowledge*. There are several concepts that figure prominently in traditional theories of knowledge. Traditional epistemologists have argued that knowledge is caused by the facts and that it is sensitive. It has also been argued that knowledge is not just the result of luck, that knowledge is safe, and that knowledge is absent in Gettier cases.⁴⁶ These same concepts figure prominently in theories of what legal proof requires. We have already examined causal and sensitivity accounts. In addition, THOMSON 1986 argues that legal proof requires something that is *not the result of luck*:

[I]t is required of a jury that it not impose liability unless it has, not merely good reason, but reason of a kind which would make it not be just luck for the jury if its verdict is true. (214)

PRITCHARD 2015 argues that proof requires *safety*:

[A]n adequacy condition on the total evidence presented at trial in support of the defendant's guilt ought to be such that it satisfies the epistemic anti-risk condition. . . i.e., it is not an easy possibility that one's belief could have been false. (457)

46. All five of these features of knowledge routinely figure in introductory epistemology texts. For example, see DANCY 1985, NAGEL 2014, GOLDMAN & MCGRATH 2015, and ICHIKAWA & STEUP 2018.

PARDO 2010 argues that proof requires something that is *absent in Gettier cases*:

[A]n appropriate connection among (1) fact finders' conclusions (however conceived), (2) the epistemic support they require, and (3) their truth, matters for legal proof. This connection arises in non-Gettier cases and is missing in Gettier cases... the goal of legal proof is *non-Gettier-ized* true and justified conclusions. (55, 57)

The common thread running throughout this literature speaks for itself. An elegant explanation of these several observations is that legal proof requires *knowledge*.

There is just one hitch. A small handful of scholars have endorsed the idea that the criminal standard of proof requires knowledge. But it has proven difficult to extend this idea to other standards of proof, such as proof by a preponderance of the evidence. REDMAYNE 2008 describes this difficulty in detail. Redmayne agrees with Thomson that the strong parallels between examples like *Lottery* and *Prisoners* suggest that legal proof requires knowledge. But he goes on to object that this solution to the problem of statistical evidence cannot be extended to civil cases:⁴⁷

There is an obvious problem with this view, however. It is plausible that whatever prevents a liability verdict in *Prisoners* also prevents a liability verdict in *Blue Bus*. If *Prisoners* is explained by a knowledge requirement for proof, then *Blue Bus* is too. But that would involve arguing that civil as well as criminal verdicts require knowledge, and that is not easy to accept. Civil verdicts require no more than proof on the balance of probabilities. This standard seems too low to satisfy the degree of justification required for knowledge. (299)

As Redmayne sees it, the civil standard of proof cannot require knowledge, since the subjective component of this standard is merely that the factfinder has greater than .5 credence that the defendant is liable. By contrast, knowing that a defendant is liable requires the factfinder to fully believe that the defendant is liable, which is a much stricter constraint. A knowledge account of the criminal standard is plausible, since outright belief in guilt is a viable constraint on conviction. But a knowledge account of the civil standard is out.⁴⁸

How to respond? Redmayne is correct that proof of liability by a preponderance of the evidence cannot require that the factfinder know *that the defendant is liable*.

47. The *Blue Bus* case mentioned here by Redmayne is a hypothetical civil case modeled after *Smith v. Rapid Transit* 58 N.E.2d 754 (1945). A car is negligently run off the road by a blue bus. The driver of the car can't identify the exact bus that caused the accident, but she can prove that the Blue Bus Company operates 60 percent of the blue buses in town, while another company operates only the remaining 40 percent. The evidence is insufficient to sustain a verdict. For an early discussion of the implications of the *Blue Bus* case for theories of statistical evidence, see TRIBE 1971.

48. For similar objections, see STEIN 2005 and BLOME-ILLMANN 2017. Although this objection is not a problem for my account, it is a problem for the knowledge account defended in PARDO 2010. Pardo applies his account to civil cases like *Blue Bus*, arguing that a verdict for the plaintiff requires knowing that the defendant is liable for the accident (52).

But there is another option available—namely, that proof by a preponderance of the evidence requires that the factfinder have knowledge of a significantly weaker content. Here is the thesis I want to defend: proof of liability by a preponderance of the evidence requires that the factfinder know that the defendant is *probably* liable. What does it mean to know merely that the defendant is probably liable—or equivalently, that the defendant is *at least .5 likely* to be liable—as opposed to simply knowing that the defendant is liable? There is a rich literature on the formal semantics of knowledge ascriptions embedding probability operators. According to my preferred interpretation, these statements are used to ascribe *probabilistic knowledge* in the sense of Moss 2018b. Rather than ascribing an outright belief to the factfinder, they say that the factfinder has certain credences and that those credences constitute knowledge.⁴⁹

This interpretation of my thesis requires an epistemological assumption—namely, that credences, just like full beliefs, are among the kinds of attitudes that can constitute knowledge. Just as you can rule out relevant alternatives to your full beliefs, you can rule out relevant alternatives to your probabilistic beliefs. Rule out enough of them, and you can acquire knowledge. For example, suppose you have a sore throat, and you have greater than .5 credence that it is caused by a strep virus, and you have less than .5 credence that you are contagious. A doctor runs some tests and arrives at just the same conclusions. After talking with the doctor, you may come to *know* that you probably have strep throat and that you are probably not contagious. But before ruling out the possibility that her test results would contradict your hunches, your probabilistic beliefs did not yet constitute knowledge.

Returning to a legal context, consider the civil standard of proof by a preponderance of the evidence. As discussed in §3.1, the defendant is proven liable by this standard only if the judge or jury has greater than .5 credence that he is liable. The knowledge account adds an objective element to this condition: the defendant is proven liable by a preponderance of the evidence if and only if this probabilistic belief constitutes knowledge. Suppose you are the factfinder in *Gatecrasher*. The plaintiff has proved that most people at the rodeo climbed over the fence. But the defendant insists that he is not just any arbitrary person at the rodeo. In effect, the defendant is raising a certain possibility—namely, that he is an individual, not represented by features of the group to which he belongs. Given the lottery-like similarity of all the possible defendants in the *Gatecrasher* scenario, this possibility is impossible to ignore. In the absence of other evidence, you may be *justified* in having .501 credence that the

49. To be more precise, these ascriptions say that the factfinder has a *probabilistic belief* that constitutes knowledge. A probabilistic belief can be a precise credence, such as .5 credence that a certain coin landed heads, or it can correspond to a range of credences, as with the belief that it is more than .5 likely that the coin landed heads. For further introductory discussion of probabilistic beliefs, see chapter 1 of Moss 2018b.

defendant is liable, and even justified in betting that he is liable. But your justified credence does not constitute knowledge. You do not know that this particular defendant is probably liable, because you can't rule out a relevant possibility that is inconsistent with this content—namely, that the defendant is an individual whose character makes him far less likely to trespass than just some arbitrary person at the rodeo. According to my account of the civil standard of proof, it follows that the plaintiff has failed to prove that the defendant is liable by a preponderance of the evidence.

The knowledge account similarly extends to other legal standards of proof. As discussed in §3.1, each standard of proof corresponds to a distinct threshold of credence. A standard of proof is met just in case the factfinder has at least that much credence in a proposition, and this probabilistic belief constitutes knowledge. Apropos of this thesis, it is worth noting that in elucidating different standards of proof, Judge Weinstein does not merely say that they require different levels of confidence, but different *levels of knowledge*:

The problem, then, is to determine what level of knowledge by the Marshal is necessary to permit this intrusion on the person. Implied in terms such as “probable cause” or “reasonableness” is a continuum of probability that the subject has been, is, or is about to be, engaged in criminal activity; it begins with no evidence of such conduct and extends to almost certainty.

To give one more example, proof of liability by clear and convincing evidence requires the factfinder to know that it is *highly* probable that the defendant is liable. An informal study conducted by Judge Weinstein found that judges in the Eastern District of New York associated this standard with probability thresholds ranging from .6 to .75, for instance, and a subsequent study by McCauliff 1982 found that hundreds of federal judges and United States Supreme Court justices associated this standard with an average probability threshold of .75.⁵⁰ Compared with the preponderance standard, proof of liability by clear and convincing evidence requires the factfinder to have an even higher credence in liability, and also that this higher credence constitutes knowledge.

3.5 Accounting for exceptional cases

I have defended an account of legal proof that is based on knowledge and also solves the general problem of statistical evidence presented by all standards of proof. So far, I have focused on the traditional problem of statistical evidence—namely, explaining the fact that statistical evidence is generally insufficient for proof. Now we can turn our attention to two additional facts worth explaining.

50. For further discussion of the former study, see *United States v. Fatico*, 458 F. Supp. 388, 410 (1978).

The first fact is that our intuitions about the sufficiency of merely statistical evidence vary greatly from case to case. It seems unconscionable to convict the defendant in *Prisoners*. But in other cases, courts have been willing to treat statistical evidence as sufficient for legal proof, and it is not at all obvious that they shouldn't. For example, in *Kramer v. Weedhopper*, Kramer was injured in a plane crash which resulted from a defective bolt sold as part of a Weedhopper airplane kit. Weedhopper purchased 90% of its bolts from Lawrence and 10% from Hughes. Kramer filed a complaint alleging strict product liability against Lawrence. The circuit court granted summary judgment to Lawrence, arguing that the available statistical evidence was insufficient to prove that the bolt that injured Kramer probably came from Lawrence. But an Illinois appellate court reversed the judgment, stating that "when two suppliers of an allegedly dangerous product exist, under circumstances which show that a defendant supplied 90% of the parts used, this evidence is sufficient to withstand a motion for summary judgment."⁵¹

Statistical evidence is also sometimes considered sufficient for a verdict in toxic tort cases where epidemiological evidence is used to prove specific causation claims. For example, in *Manko v. United States*, Manko developed Guillain-Barré syndrome within weeks of receiving a swine flu vaccination. He alleged that the latter had caused the former. An expert testified that the relative risk of contracting the syndrome after vaccination was greater than 2, meaning that the syndrome was more than twice as prevalent in people who had received the vaccine as compared with those who hadn't. Based on this evidence, the court entered a judgment for the plaintiff, stating, "Because the relative risk of contracting GBS during the period 11–16 weeks after vaccination is greater than 2, it is more likely than not that plaintiff's swine flu vaccination caused his GBS."⁵² The use of relative risk to prove specific causation is controversial, to say the least.⁵³ But whether or not these judgments are correct, it is worth noting that they do not offend our intuitions as much as traditional cases of verdicts based on statistical evidence.

This variation in our intuitions about statistical evidence is unexpected on causal and sensitivity accounts of legal proof. The statistical facts used in *Weedhopper* and *Manko* are no more causally connected to individual claims in these cases than they are in *Gatecrasher*, nor are they more sensitive to the specific facts of causation that they are used to prove. By contrast, varying intuitions about statistical evidence are to be expected on the knowledge account. On my account, statistical evidence suffices

51. *Kramer v. Weedhopper of Utah, Inc.*, 141 Ill. App. 3d 217, 220 (1986). Another similar case often cited in the literature is *Kaminsky v. Hertz Corp.*, 288 N.W.2d 426 (Mich. App. 1979).

52. *Manko v. United States*, 636 F. Supp. 1434 (W.D. Mo. 1986).

53. See HAACK 2014 for a detailed history of this practice, as well as several compelling arguments that relative risk factors are neither necessary nor sufficient for proof of specific causation claims.

to prove causation just in case the factfinder knows that causation is more than .5 likely. Put another way, the factfinder must rule out all relevant possibilities according to which causation is no more than .5 likely. Of course, this raises the central question of just what counts as a relevant possibility. As discussed in §2.2, the answer to this question is highly context sensitive. We have seen that the relevance of a possibility may depend on its lottery-like structure. In addition, it is often held that whether a possibility is relevant depends partly on what is at stake—for instance, the cost of having a false belief if the possibility were to obtain.⁵⁴ Falsely profiling an individual person as having a negative character trait might be morally different from other instances of false profiling, such as falsely profiling inanimate objects or corporate defendants. False profiling that harms some specific person—as in *Dog Bite*, for instance—might also carry distinctive moral costs.⁵⁵ By contrast, false verdicts in some product liability and toxic tort cases may lack any similar costs. To sum up, differences in the stakes of a false verdict may naturally lead to some variation in how easy it is for jurors to have knowledge—which, on my account, explains the variation in our judgments about what is sufficient for legal proof.

This discussion of the moral stakes of profiling brings us to a second fact to be explained. Some statistical evidence is not only insufficient for proof, but inadmissible at trial. As *TRIBE 1971* reminds us, these are very different claims:

[T]he fact that mathematical evidence taken alone can rarely, if ever, establish the crucial proposition with sufficient certitude to meet the applicable standard of proof does not imply that such evidence—when properly combined with other, more conventional, evidence in the same case—cannot supply a useful link in the process of proof. (1350)

Hence we need some additional explanation for why courts sometimes exclude statistical evidence as irrelevant, unable to serve even as a link in the process of proof. For example, statistical facts about crime rates among residents of a given neighborhood are irrelevant when it comes to proving the criminal behavior of a particular resident of that neighborhood. Similarly, the fact that most assaults in a given neighborhood involve illegal firearms is irrelevant when it comes to proving that some particular assault in that neighborhood involved an illegal firearm. On the basis of an examination of hundreds of state and federal cases, *KOEHLER 2002* observes that “[c]ourts are likely to find base rates irrelevant when they smack of guilt by trait association” (383).

54. Compare the *Rule of High Stakes* from *LEWIS 1996*: “when error would be especially disastrous, few possibilities are properly ignored” (556). See *DEROSE 1992* for a classic discussion of stakes and knowledge ascriptions, though see also *WORSNIP 2015* and *ANDERSON & HAWTHORNE 2019* for important arguments that it is difficult to spell out the notion of stakes that is central to this literature.

55. See *MOSS 2018a* for a detailed discussion of how the moral stakes of a belief may have an epistemic impact on whether that belief constitutes knowledge.

Again, it is hard for other epistemic accounts to explain these observations. Causal and sensitivity accounts merely impose a necessary condition on the sufficiency of evidence, and it is not obvious how to extend these accounts to provide an additional necessary condition for admissibility. By contrast, the knowledge account can explain why some statistical inferences are impermissible. Statistical evidence might justify the factfinder in having a higher credence in a claim. But whether this higher credence constitutes knowledge depends on whether the factfinder can rule out certain possibilities, and some possibilities may be epistemically resilient, such that no evidence could possibly rule them out. The possibility of being a brain in a vat is epistemically resilient in the context of an epistemology classroom, designed in order to be irrefutable by observation or reflection. In the context of a criminal trial, other possibilities are resilient, including the possibility that the defendant has individual traits that would undermine any alleged probative value of statistics having to do with racial categories. As with brain-in-a-vat hypotheses in a classroom context, there may be features of the courtroom context that prevent jurors from ruling out this possibility, thereby making it impossible for certain statistical facts to make even a minor contribution to an inference that could provide them with the knowledge they need to convict. Jurors are not morally compelled to assume that certain statistical inferences *are in fact* misleading, but only that they *might be* misleading—yet as long as this possibility is taken for granted, those same inferences cannot help ground knowledge.

A consistent theme has emerged from our discussion—namely, that moral considerations can play an important epistemic role in legal contexts. As a result of this fact, my knowledge account incorporates a lot of the desirable features of moral and epistemic accounts of legal proof. Although the knowledge account is itself epistemic, moral considerations play an important role in it, since there may be moral reasons why some possibilities of error involve high stakes, or can be safely ignored by jurors, or are especially resilient against further evidence. This avenue of influence for moral concerns underscores a broader theme of this paper—namely, that legal proof is not merely a matter of justifying a subjective probability judgment. The epistemic relevance of a possibility does not depend only on its probability, but also on its moral features. Accordingly, on the knowledge account, moral features of statistical inferences can make a difference to their proper role in legal proof.

4 Responses to objections

Having set out some positive arguments for my account, let me now address two important objections. The first objection, which is perhaps the prevailing objection in

philosophical discussions of legal proof, is that the knowledge account delivers factive standards of legal proof. According my account, an innocent defendant cannot be proven guilty.⁵⁶ Some find this result unacceptable. As GARDINER 2019 puts it,

[I]t is a desideratum of an account of legal evidence that—if evidence is compelling but misleading—the burden can be satisfied even if the judgment is false...an innocent defendant can appear guilty beyond reasonable doubt. (7)

Similarly, BLOME-TILLMANN 2015 argues:

A final problem to be addressed here concerns wrongful convictions...In some such cases of wrongful conviction, the defendant is found liable on the basis of very strong but ultimately misleading evidence. In such cases, the court is not at fault. The standards of proof have been met by strong, if misleading, evidence. Call such cases *no-fault wrongful convictions*.

Blome-Tillmann goes on to argue that no-fault wrongful convictions are impossible according to factive standards of proof, and he rejects factive standards for this reason. It is easy to appreciate this worry. Suppose that the prosecution has presented such a strong case that the only remaining possibilities of innocence involve far-fetched conspiracy theories. And now suppose that as it happens, the defendant *is* the innocent victim of a conspiracy. An opponent of the knowledge account might say, “In situations like these, don’t we *want* the jury to set aside the far-fetched conspiracy theory and find the defendant guilty?”⁵⁷

The correct response is, “Well, we do and we don’t.” There are multiple dimensions along which we can evaluate jurors as they deliberate. There is one sense in which the jury should convict the defendant in the imagined case. But there is also a sense in which they shouldn’t convict the innocent defendant—because there is a sense in which there is clearly something bad about false convictions, whether or not they result from misleading evidence. It is this second sort of intuitive judgment that is at work in the observation by WILLIAMS 1980 that “There is a miscarriage of justice whenever an innocent man is convicted” (104), and also at work in the eighteenth-century judgment that convicting an innocent defendant was a mortal sin. In short, there is something good about convicting an innocent person on misleading evidence and also something bad about it. Any decent account of legal proof should offer explanations of both judgments.

56. For simplicity, I focus on the criminal standard in this section. The civil standard of proof by a preponderance of the evidence is also factive, in the sense that a defendant cannot be proved probably liable unless the defendant is probably liable. For a detailed discussion of the notion of factivity as it applies to probabilistic knowledge, see chapter 5 of MOSS 2018b.

57. For additional discussion of this objection to the knowledge account, see §2B of BELTRÁN 2006 and fn. 32 of BLOME-TILLMANN 2017.

The knowledge account of legal proof delivers both of these explanations. Suppose the jurors convict the innocent victim of the secret government conspiracy. Although they convict this defendant without proof that he is guilty, they are entirely blameless for doing so, since they are doing their subjective best to satisfy the objective legal rule they are following. They are subjectively in the right, despite being objectively in the wrong. This combination of normative judgments is familiar, as it can be found wherever we find externalist norms—which is to say, everywhere. Consider these ordinary norms:

- (3) Turn off the stove as soon as the water starts boiling.
- (4) Water the garden if and only if it doesn't rain.
- (5) Set the Passover table for however many guests are coming, plus one.
- (6) Call the tennis ball out if and only if it falls completely outside the line.
- (7) Hang one light in the tower if they are coming by land, and two if by sea.
- (8) Give your money to whichever charity has the lowest overhead.

You can have a justified belief that you are following these norms, even when you aren't actually following them. In that case, you are not doing the objectively right thing. But you may be blameless, and even praiseworthy, for doing what you believe is the objectively right thing to do. The same goes for the knowledge norm of conviction:

- (9) Convict the defendant if and only if you know he is guilty.

For every primary objective norm that tells you what you ought to do, there are secondary subjective norms, such as the instruction to do whatever you believe will result in the satisfaction of the primary norm. As far as the normative landscape goes, the subjective norms are grounded in the objective norms from which they are predictably derived.⁵⁸ Recall Gardiner's objection that "an innocent defendant can appear guilty beyond reasonable doubt" (7). Gardiner is right. But notice what Gardiner does not say—namely, that an innocent defendant can be *proven* guilty beyond a reasonable doubt. In some cases of misleading evidence, an innocent defendant appears to have been proven guilty. The criminal standard of proof appears to be satisfied. This is sufficient to explain our intuition that there is a sense in which the jurors should convict, and that is enough to resolve the first objection to the knowledge account.

The second objection to my account is a broader worry about its externalist char-

58. The same response also applies to internalist objections to other knowledge norms, such as knowledge norms of belief and assertion. For further discussion, see LASONEN-AARNIO 2010, HAWTHORNE & SRINIVASAN 2013, and WILLIAMSON 2015.

acter. At first glance, it might seem unsettling that whether a jury ought to convict a defendant could depend on whether their evidence is misleading, when they have no way of determining whether that's true. SMITH 2018 raises something like this internalist objection in the following passage:

If two courts were presented with equivalent bodies of evidence against two individuals charged with equivalent crimes, could it really be acceptable for them to reach different verdicts—for one individual to be found guilty and the other innocent—even if there was some variation in external circumstances? (1205)

As stated, this objection is not hard to answer. In some cases, it might turn out that the evidence is *permissive*, in the sense that reasonable juries could disagree about what verdict it supports. In such cases, it should be acceptable for reasonable juries to reach different verdicts. But setting aside cases of permissive evidence, we can formulate an objection that is similar in spirit but harder to answer:

If two courts are presented with equivalent bodies of evidence against two individuals charged with equivalent crimes, and the juries form exactly the same beliefs, and in fact the juries are *exact intrinsic duplicates*, could the juries be *required* to reach different verdicts in virtue of some variation in external circumstances?

The knowledge account says that this is possible. But at first glance, it might seem inappropriate to hold juries to any standard of this sort, since the inscrutability of the standard necessarily entails that juries will not be able to control whether they are following it.

Here again, it is helpful to reflect on precedents provided by ordinary norms. In particular, it is helpful to think about legal statutes. Consider the following example involving statutes violated by drunk drivers:

Drunk Drivers: Two drunk drivers pull out onto the highway. They are exact intrinsic duplicates, both with the same blurry field of vision and delayed reflexes. Both swerve unexpectedly into oncoming traffic. The unlucky drunk driver causes a fatal collision and is convicted of vehicular homicide. The lucky driver causes a minor collision and is convicted of negligent driving.

The drivers have no control over the circumstances that determine which offense they commit. The relevant offenses concern not just internal states, but also their external effects.⁵⁹ Legally speaking, it is not only important to avoid risky actions, but also important to avoid killing people as a result of undertaking risky actions. In short, many ordinary legal statutes are externalist norms, norms that give rise to legal luck.

59. Arguably, the same sort of luck adheres to moral judgments, though of course that's more controversial. For a radical discussion of legal and moral luck and the relationship between them, see ENOCH 2018.

We have arrived at an externalist thesis about legal statutes governing drivers on the highway. What about legal standards governing jurors in a courtroom? In fact, the dialectic is remarkably similar. Consider the following pair of cases:

Lucky Lucy: Lucy is a juror in a criminal trial. She is not epistemically perfect. For instance, when she listens to people talk for a long time, she occasionally gets sleepy and, without realizing it, misses gathering some information. Despite her imperfections, Lucy comes to know that the defendant is guilty after listening to a lot of compelling evidence presented by the prosecution at trial.

As it happens, Lucy could have easily failed to have this knowledge.⁶⁰ Lucy knows that the defendant is guilty only because she has done something epistemically blameworthy. At one point in the trial, she happened to miss hearing a false statement made by a friend of the defendant. If she had been paying closer attention, Lucy would have been successfully tricked by the false testimony into believing that the defendant might be innocent.

Lucy is lucky. The evidence that she missed was misleading, and her epistemic imperfections have not prevented her from gaining knowledge. Since she knows that the defendant is guilty, she can correctly vote to convict. Unfortunately, her duplicate is not so lucky:

Unlucky Ursula: Ursula is an intrinsic duplicate of Lucy, serving as a juror at a duplicate trial. Like Lucy, Ursula hears a lot of compelling evidence against the defendant, and on the basis of this evidence, she forms a justified belief that the defendant is guilty. She also misses hearing a brief statement made by a friend of the defendant, which would have caused her to believe that the defendant might be innocent.

Unfortunately, in Ursula's case, the defendant is actually innocent. The prosecution's evidence is misleading, and the statement made by the friend of the defendant is genuinely exculpatory. Ursula is justified in believing that the defendant is guilty, but her belief is false, and it isn't knowledge. If she had just been paying closer attention, Ursula wouldn't have voted to convict an innocent person.

Lucky Lucy and *Unlucky Ursula* elicit just the same intuitive judgments as traditional illustrations of legal and moral luck. An imperfect juror who falsely convicts an innocent person is falling short of the ideal set by our legal standards. Unlucky Ursula deserves some blame for the false conviction that she helped cause, where Lucky Lucy is not blameworthy in this way. In order for there to be any such normative difference between intrinsically duplicate jurors, there must be at least some externalist legal standards governing their verdicts. The knowledge account succeeds in providing standards of proof of just this sort.

60. To forestall confusion: my claim here—that Lucy could have easily failed to believe what she actually knows—is compatible with the claim that her knowledge is *safe*, i.e. that she couldn't have easily falsely believed that the defendant was guilty.

In defending the knowledge account, we have come across a novel variety of legal luck. In addition to first-order legal statutes that are sensitive to luck, our standards of proof are also sensitive to luck. And given that these standards are sensitive to luck, the second objection to the knowledge account fails. To be clear, it is not the case that the second objection fails *only if* our legal standards are luck-sensitive. It might turn out that there are externalist standards of proof but that there is no legal luck, i.e. that whether a juror is blameworthy for violating externalist legal standards depends only on internal characteristics that are under her control. In other words, I have defended a claim that is sufficient but not necessary to demonstrate that there are externalist standards of proof. Happily, my defensive argument also serves as a final positive argument for my account. Any account of proof should explain the fact that jurors with justified beliefs but epistemic shortcomings may be distinctively blameworthy in cases where their shortcomings lead to false convictions, and the knowledge account successfully explains this fact.

References

- ANDERSON, CHARITY & JOHN HAWTHORNE. 2019. "Knowledge, Practical Adequacy, and Stakes." Ms., Department of Philosophy, Baylor University and University of Southern California. Forthcoming in *Oxford Studies in Epistemology*, vol. 6.
- BELTRÁN, JORDI FERRER. 2006. "Legal Proof and Fact Finders' Beliefs." *Legal Theory*, vol. 12 (4): 293–314.
- BENTHAM, JEREMY. 1843. *The Works of Jeremy Bentham, vol. 6 (The Rationale of Judicial Evidence)*. William Tait, Edinburgh. Edited by John Bowring.
- BLOME-TILLMANN, MICHAEL. 2015. "Sensitivity, Causality, and Statistical Evidence in Courts of Law." *Thought*, vol. 4 (2): 102–12.
- . 2017. "'More Likely Than Not'—Knowledge First and the Role of Bare Statistical Evidence in Courts of Law." In *Knowledge First: Approaches in Epistemology and Mind*, ADAM CARTER, EMMA GORDON & BENJAMIN JARVIS, editors, 278–92. Oxford University Press, Oxford.
- COHEN, L. JONATHAN. 1977. *The Probable and the Provable*. Clarendon, Oxford.
- . 1981. "Subjective Probability and the Paradox of the Gatecrasher." *Arizona State Law Journal*, vol. 2 (2): 627–34.
- . 1987. "On Analyzing the Standards of Forensic Evidence: A Reply to Schoeman." *Philosophy of Science*, vol. 54 (1): 92–97.
- DANCY, JONATHAN. 1985. *Introduction to Contemporary Epistemology*. Blackwell, Oxford.
- DEROSE, KEITH. 1992. "Contextualism and Knowledge Attributions." *Philosophy and Phenomenological Research*, vol. 52 (4): 913–29.
- . 2002. "Assertion, Knowledge, and Context." *Philosophical Review*, vol. 111 (2): 167–203.
- DUFF, ANTONY, LINDSAY FARMER, SANDRA MARSHALL & VICTOR TADROS. 2007. *The Trial on Trial (vol. 3): Towards a Normative Theory of the Criminal Trial*. Hart, Oxford.
- ENOCH, DAVID. 2018. "Luck Between Morality, Law, and Justice." *Theoretical Inquiries in Law*, vol. 9 (1): 23–59.
- ENOCH, DAVID, LEVI SPECTRE & TALIA FISHER. 2012. "Statistical Evidence, Sensitivity, and the Legal Value of Knowledge." *Philosophy and Public Affairs*, vol. 40 (3): 197–224.

- FOGELIN, ROBERT. 2000. "Contextualism and Externalism: Trading in One Form of Skepticism for Another." *Philosophical Issues*, vol. 10 (1): 43–57.
- GARDINER, GEORGI. 2019. "Legal Burdens of Proof and Statistical Evidence." Ms., Department of Philosophy, Oxford University. Forthcoming in *Routledge Handbook of Applied Epistemology*, eds. David Coady and James Chase. Routledge.
- GOLDMAN, ALVIN & MATTHEW MCGRATH. 2015. *Epistemology: A Contemporary Introduction*. Oxford University Press, New York.
- GREEN, THOMAS. 1985. *Verdict According to Conscience: Perspectives on the English Criminal Trial Jury, 1200–1800*. Univ. of Chicago Press, Chicago.
- HAACK, SUSAN. 2014. "Risky Business: Statistical Proof of Specific Causation." In *Evidence Matters: Science, Proof, and Truth in the Law*, 264–93. Cambridge University Press, Cambridge.
- HARMAN, GILBERT. 1986. *Change in View: Principles of Reasoning*. MIT Press, Cambridge.
- HAWLES, JOHN. 1680. *The Englishman's Right: A Dialogue between a Barrister at Law, and a Juryman*. Printed for Richard Janeway, in Queenshead Alley in Paternoster-Row, London. Reprinted New York: Garland, 1978.
- HAWTHORNE, JOHN. 2004. *Knowledge and Lotteries*. Oxford University Press, Oxford.
- HAWTHORNE, JOHN & AMIA SRINIVASAN. 2013. "Disagreement Without Transparency: Some Bleak Thoughts." In *The Epistemology of Disagreement: New Essays*, DAVID CHRISTENSEN & JENNIFER LACKEY, editors, 9–30. Oxford University Press, Oxford.
- H.E. 1664. *The Juryman charged; or a Letter to a Citizen of London*. n.p., London.
- HO, HOCK LAI. 2008. "Book Review: The Origins of Reasonable Doubt—Theological Roots of the Criminal Trial by James Q. Whitman." *Singapore Journal of Legal Studies*, vol. 2008 (2): 455–62.
- ICHIKAWA, JONATHAN & MATTHIAS STEUP. 2018. "The Analysis of Knowledge." In *The Stanford Encyclopedia of Philosophy (Summer 2018 Edition)*, EDWARD N. ZALTA, editor. Available at <https://plato.stanford.edu/archives/sum2018/entries/knowledge-analysis/>.
- KOEHLER, JONATHAN J. 2002. "When Do Courts Think Base Rate Statistics are Relevant?" *Jurimetrics*, vol. 42 (4): 373–402.

- LASONEN-AARNIO, MARIA. 2010. "Unreasonable Knowledge." *Philosophical Perspectives*, vol. 24 (1): 1–21.
- LAUDAN, LARRY & HARRY SAUNDERS. 2009. "Re-Thinking the Criminal Standard of Proof: Seeking Consensus about the Utilities of Trial Outcomes." *International Commentary on Evidence*, vol. 7 (2): 1–34.
- LEWIS, DAVID K. 1996. "Elusive Knowledge." *Australasian Journal of Philosophy*, vol. 74 (4): 549–67.
- LILLQUIST, ERIK. 2002. "Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability." *U.C. Davis Law Review*, vol. 36 (1): 85–198.
- LITTLEJOHN, CLAYTON. 2017. "Truth, Knowledge, and the Standard of Proof in Criminal Law." Ms., Department of Philosophy, King's College London. Forthcoming in *Synthese*.
- MAGNUSSEN, SVEIN, DAG ERIK ELIERTSEN, KARL HALVOR TEIGEN & ELLEN WESSEL. 2014. "The Probability of Guilt in Criminal Cases: Are People Aware of Being 'Beyond Reasonable Doubt'?" *Applied Cognitive Psychology*, vol. 28 (2): 196–203.
- MCCAULIFF, C. M. A. 1982. "Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees?" *Vanderbilt Law Review*, vol. 35 (6): 1293–1336.
- MOORE, G. E. 1962. *Commonplace Book: 1919–1953*. George Allen & Unwin Ltd., London.
- MOSS, SARAH. 2018a. "Moral Encroachment." *Proceedings of the Aristotelian Society*, vol. 118 (2): 177–205.
- . 2018b. *Probabilistic Knowledge*. Oxford University Press, Oxford.
- NAGEL, JENNIFER. 2014. *Knowledge: A Very Short Introduction*. Oxford University Press, Oxford.
- NESSON, CHARLES. 1979. "Reasonable Doubt and Permissive Inferences: The Value of Complexity." *Harvard Law Review*, vol. 92 (6): 1187–1225.
- NOZICK, ROBERT. 1981. *Philosophical Explanations*. Harvard University Press, Cambridge.
- PALEY, WILLIAM. 1785. *Principles of Moral and Political Philosophy*. Printed for R. Faulder, London.

- PARDO, MICHAEL. 2005. "The Field of Evidence and the Field of Knowledge." *Law and Philosophy*, vol. 24 (4): 321–92.
- . 2010. "The Gettier Problem and Legal Proof." *Legal Theory*, vol. 16 (1): 37–57.
- PRITCHARD, DUNCAN. 2015. "Risk." *Metaphilosophy*, vol. 46 (3): 436–61.
- REDMAYNE, MIKE. 2008. "Exploring the Proof Paradoxes." *Legal Theory*, vol. 14 (4): 281–309.
- ROTH, ANDREA. 2010. "Safety in Numbers? Deciding When DNA Alone is Enough to Convict." *New York University Law Review*, vol. 85 (4): 1130–85.
- SAKS, MICHAEL J. & ROBERT F. KIDD. 1980. "Human Information Processing and Adjudication: Trial by Heuristics." *Law & Society Review*, vol. 15 (1): 123–60.
- SAND, LEONARD & 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*, vol. 78 (3): 1359–75.
- SCHAUER, FREDERICK. 2003. *Profiles, Probabilities, and Stereotypes*. Harvard University Press, Cambridge.
- SHAVIRO, DANIEL. 1989. "Statistical-Probability Evidence and the Appearance of Justice." *Harvard Law Review*, vol. 103 (2): 530–54.
- SMITH, MARTIN. 2018. "When Does Evidence Suffice for Conviction?" *Mind*, vol. 127 (508): 1193–1218.
- SORENSEN, ROY. 2006. "Future Law: Prepunishment and the Causal Theory of Verdicts." *Notus*, vol. 40 (1): 166–83.
- SOSA, ERNEST. 1999. "How to Defeat Opposition to Moore." *Philosophical Perspectives*, vol. 13: 141–53.
- STANLEY, JASON. 2005. *Knowledge and Practical Interests*. Oxford University Press, Oxford.
- STEEL, SANDY. 2015. "Justifying Exceptions to Proof of Causation in Tort Law." vol. 78 (5): 729–58.
- STEIN, ALEX. 2005. *Foundations of Evidence Law*. Oxford University Press, Oxford.
- THOMSON, JUDITH JARVIS. 1986. "Liability and Individualized Evidence." *Law and Contemporary Problems*, vol. 49 (3): 199–219.

- TRIBE, LAURENCE H. 1971. "Trial by Mathematics: Precision and Ritual in the Legal Process." *Harvard Law Review*, vol. 84 (6): 1329–93.
- UNGER, PETER. 1975. *Ignorance: A Case for Scepticism*. Clarendon Press, Oxford.
- VOGEL, JONATHAN. 1990. "Are There Counterexamples to the Closure Principle?" In *Doubling: Contemporary Perspectives on Scepticism*, MICHAEL D. ROTH & GLENN ROSS, editors, 13–25. Kluwer, Dordrecht.
- WASSERMAN, DAVID T. 1991. "The Morality of Statistical Proof and the Risk of Mistaken Liability." *Cardozo Law Review*, vol. 13 (2–3): 935–76.
- WHITMAN, JAMES. 2008. *The Origins of Reasonable Doubt: Theological Roots of the Criminal Trial*. Yale University Press, New Haven.
- WIGMORE, JOHN HENRY. 1923. *A Treatise on the Anglo-American System of Evidence in Trials at Common Law*, 3rd edn. Little, Brown and Company, Boston.
- WILLIAMS, GLANVILLE. 1980. "A Short Rejoinder." *Criminal Law Review*, 403–7.
- WILLIAMSON, TIMOTHY. 2000. *Knowledge and its Limits*. Oxford University Press, Oxford.
- . 2015. "Justifications, Excuses, and Sceptical Scenarios." Ms., Department of Philosophy, Oxford University. Forthcoming in *The New Evil Demon: New Essays on Knowledge, Rationality and Justification*, eds. Julien Dutant and Daniel Dorsch. Oxford University Press, Oxford.
- WORSNIP, ALEX. 2015. "Two Kinds of Stakes." *Pacific Philosophical Quarterly*, vol. 96 (3): 307–24.
- ZUCKERMAN, ADRIAN A. S. 1986. "Law, Fact or Justice?" *Boston University Law Review*, vol. 66 (4): 487–508.