CrossMark

Truth, knowledge, and the standard of proof in criminal law

Clayton Littlejohn[1]

Abstract Could it be right to convict and punish defendants using only statistical evidence? In this paper, I argue that it is not and explain why it would be wrong. This is difficult to do because there is a powerful argument for thinking that we *should* convict and punish defendants using statistical evidence. It looks as if the relevant cases are cases of decision under risk and it seems we know what we should do in such cases (i.e., maximize expected value). Given some standard assumptions about the values at stake, the case for convicting and punishing using statistical evidence seems solid. In trying to show where this argument goes wrong, I shall argue (against Lockeans, reliabilists, and others) that beliefs supported only by statistical evidence are epistemically defective and (against Enoch, Fisher, and Spectre) that these epistemic considerations should matter to the law. To solve the puzzle about the role of statistical evidence in the law, we need to revise some commonly held assumptions about epistemic value and defend the relevance of epistemology to this practical question.

✉ Clayton Littlejohn
clayton.littlejohn@kcl.ac.uk

1   King's College London, London, UK

1 Introduction

Let's consider an example:

> Prisoners
>
> 100 prisoners are exercising in the prison yard. Suddenly 99 of them attack the guard, putting into action a plan that the 100th prisoner knew nothing about. The 100th prisoner played no role in the assault and could have done nothing to stop it. There is no further information that we can use to settle the question of any particular prisoner's involvement (Redmayne 2008).

Knowing what we know and knowing that we'll never know more than this, would it be permissible to convict and punish a prisoner chosen at random from the yard?

This is a difficult question. While I think the following thesis is counterintuitive, there are powerful arguments for it:

> Punish: It is permissible to punish the defendant in Prisoners (and similar cases where the only evidence of guilt is statistical evidence).[1]

There is a related epistemological thesis that I think we should consider alongside Punish:

> Believe: It is permissible to believe that the defendant is guilty in Prisoners (and similar cases where the only evidence of guilt is statistical evidence).[2]

While I also think this thesis is counterintuitive, a plausible line of argument supports this thesis, too.

In this paper, I shall argue that we should not convict and punish defendants in criminal trials on the basis of statistical evidence alone. In the current debate, many of the philosophers critical of Punish appeal to intuitions that conflict with Believe. Ultimately, I think that we can build a compelling case against both theses and that the best argument against Punish is built on an argument against Believe. In building this case against Believe, we face two significant obstacles. First, the most powerful argument for Punish, in my view, suggests that the permissibility of punishment has little to do with the epistemological considerations operative in debates about Punish. We need to show that epistemological debates about statistical evidence and the justification of full belief should matter to the law. Second, even if we could show that these epistemological considerations are relevant to the debates about the permissibility of

---

[1] The disagreement we'll focus on isn't about the justification of punishment, per se, but about the justification of punishment based on statistical evidence. For defenses of Punish, see Lempert (1977) and Steele (MS). For arguments against the use of statistical evidence in civil and criminal law, see Apt (2016), Blome-Tillman (2015, MS), Buchak (2013), Colyvan et al. (2001), Enoch et al. (2012), Lo (2008), Moss (forthcoming), Pritchard (forthcoming b), Redmayne (2008), Smith (2016, forthcoming), Staffel (forthcoming), Thomson (1986), and Tribe (1971). This paper focuses on criminal law.

[2] Comesana (2009) offers a partial defense of the view that we can have justified (full) beliefs based on statistical evidence. Sosa (2015) does, too. See also Dorst (forthcoming), Easwaran (2016), Foley (2009), and Sturgeon (2008). For arguments that such evidence cannot justify full belief, see Adler (2002), Kelp (forthcoming), Littlejohn (2012), Nelkin (2000), Ryan (1991), Smith (2016), and Williamson (2000).

punishing on the basis of statistical evidence, we need to show where the argument for Believe goes wrong. We need to show that the justification of full belief can require more than grounds that warrant a high degree of confidence.

The discussion is complicated because it touches upon controversial issues in law, moral theory, and epistemology. Let me provide a brief overview of the paper and my argument. In Sect. 2, I shall present the argument for Punish. In Sect. 3, we shall review some of the more important objections to Punish. In Sect. 4, I shall provide an argument for Believe that parallels the argument for Punish. As I see things, the case against Punish that we find in the literature suffers from two problems. The first problem is that the case is built on epistemological considerations that are supposed to show that Believe is mistaken. It is not clear what relevance these considerations have in a debate about Punish. The argument for Punish presented in Sect. 2 is valid and it's not clear which premise, if any, we should reject if we reject Believe. In short, we need to show that our two theses, Believe and Punish, are related and that an argument against the first can support an argument against the second. The second problem concerns the operative epistemological assumptions in the extant discussions of Punish. Because we have a valid argument for Believe that rests on some widely held assumptions about epistemic value, we have to do so work to show that Believe is mistaken. In short, we need to show that there is something wrong with fully believing a defendant to be guilty on the basis of statistical evidence alone and address the argument for Believe.

In Sects. 5 and 6, I explain why I think we should reject Punish and Believe. The rough idea is this. Although the justification of action does not typically depend upon the justification of any particular set of beliefs (e.g., the justification of betting on Chelsea or taking an umbrella does not require, say, a justification of believing that Chelsea will win or that it will rain this afternoon), there are some acts that can only be justified if the agent can be guided by the right kinds of reasons. Actions that serve to express blame are an example. Punishment is an example. If it would be unreasonable to take some defendant to be involved in some criminal act, for example, this has a bearing on the justification of punishment in a way that, say, the justification of taking an umbrella doesn't turn on whether we would be justified in fully believing that it will rain or sleet. As it happens, I think a case can be made for thinking that the justification of (full) beliefs concerning a defendant's guilt require more than grounds that would warrant a high degree of confidence in the defendant's guilt. In Sect. 7, I introduce a gnostic approach to epistemic value that should help us see why statistical evidence might fail to justify full belief (even when it justifies a high degree of confidence). In turn, it should help us see why considerations of expected epistemic value do not support Believe. In Sect. 8, I argue that we need this gnostic framework to solve the puzzles under consideration and discuss the significance of this for some sophisticated veritist views in epistemology inspired by Sosa's work on epistemic value.

2 An argument for punish

Our discussion concerns issues in the philosophy of law, so some stage setting is in order. I shall assume that punishment is sometimes justified and that the justification

of punishment involves the justification of the imposition of the harms reasonably expected to come with punishment. If we think that it is *prima facie* wrong to act in ways that we should expect will result in these harms, we should hope that there is some good that can come from imposing the punishment. For the sake of this discussion, let us suppose that there is some such good but not worry too much about what that good might be. (If we do *not* assume this, it will be difficult to get the puzzle off of the ground because it might seem that *all* punishment is unjustified, in which case the side that things we shouldn't punish in our example win the debate before it starts.)

We shall also assume that any just system of punishment involves some backward-looking elements. We don't have to assume that the potential goods that come from punishing people are characterized in backward-looking terms (e.g., retribution), only that we cannot impose the harms that come with punishment without any concern about whether the harms would be suffered by the guilty. While just punishment might serve a number of aims, it can only be just if, *inter alia*, it is a response to something that the defendant is answerable for.

When we think about things in these terms, we might say that the system that interests us recognizes that there is something good about punishing the guilty, something bad about punishing the innocent, that there is nothing good nor bad about failing to punish the innocent, and (most controversially, perhaps) something bad about failing to punish someone who is guilty. If these are our values (or close enough to them for the purposes of this discussion), it will help shape our understanding of what kind of evidence we'd need to rightly punish.

Let's suppose that this matrix represents the kinds of values we want our criminal justice system to serve. We'll formulate a decision-matrix with two states, two prospective acts, and assign values to the act-state pairs as follows: The matrix is described

Veritist Decision-Matrix

	Guilty	Innocent
Punish	+1	−10
Don't Punish	−1	0

as a 'veritist' matrix because it is concerned with outcomes in which a defendant is correctly or incorrectly taken to be guilty. Framed in this way, the states we should be concerned with are those that determine whether a belief in the defendant's guilt would be correct or incorrect. This assignment of values to outcomes might not capture your take on the magnitudes of the relevant goods or evils, but readers are invited to tinker with the values to see if this would change the situation.

With these values and probabilities, it's clear that Punish maximizes the relevant expected value.[3] Once we see this, we have our first argument for Punish:

---

[3] To calculate the expected value of Punish, we sum the product of the probabilities and values of <Punish, Guilty> and <Punish, Innocent>. To calculate the expected value of Don't Punish, we sum the product of the probabilities and values of <Don't Punish, Guilty> and <Don't Punish, Innocent>. When we compare the two results, we see the expected value of Punish exceeds that of Don't Punish.

Primary Argument for Punish

P1. Rationality requires us to maximize expected value.

P2. Punish is the option that maximizes expected value.

P3. So, rationality requires us to punish.

P4. It would be right to do what rationality requires.[4]

P5. So, it would be right to punish.[5]

Many of us instinctively feel that there is something wrong with Punish. The primary argument for Punish might make us worry about these intuitions. The argument is valid. We've used *our* values in formulating our decision-matrix. (Readers should remember that they are invited to tinker with the values in the matrix, but I ask them to remember that I can increase the number of prisoners to reinstate the argument.) It is hard to see how we can reject (P2). This case looks like a straightforward case of decision under risk. It seems plausible that we ought to maximize expected value in such cases. Because of this, it's hard to see how we can reject (P1) or (P4).

3 Against punish

Let's look at three objections to Punish. While these objections aren't decisive as stated, they'll help us see why we need the resources introduced later.

3.1 Tribe and reasonable doubt

Some object to the primary argument for Punish on the grounds that any conviction on the basis of statistical evidence violates the norm that juries should convict only when a defendant's guilt is beyond reasonable doubt. Tribe, for example, has proposed that the acceptance of Punish, 'could dangerously undermine …the values surrounding the notion that juries should convict only when guilt is beyond real doubt' (1971: p. 1372). He then offers these remarks:

> …to say that society recognizes the necessity of tolerating the erroneous "conviction of some innocent suspects in order to assure the confinement of a vastly

---

[4] Gibbons (2013) and Lord (forthcoming) argue that what an agent ought to do is determined by the things that make it rational or reasonable for her to act. Zimmerman (2008) argues that what an agent ought to do is determined by what would maximize a kind of expected value.

[5] It might seem that the argument is more complicated than it needs to be. I've stated it this way because there are interesting disagreements about both the relationship between rationality and obligation (e.g., between objectivists who think that there can be obligations to refrain from doing things that we're rationally required to do and authors who think that rationality and obligation are more intimately connected) and about the normative significance of expected value (e.g., between non-consequentialists who think that we should sometimes follow principles when we know that doing so stands in the way of promoting some good and those who think that we should use the tools of decision-theory to determine what to do). [] defends a version of this argument and we can find seeds of it in Lempert (1977).

larger number of guilty criminals" is not at all to say that society does, or should, embrace a policy that juries, conscious of the magnitude of their doubts in a particular case, ought to convict in the face of this acknowledged and quantified uncertainty. It is to the complex difference between these two propositions that the concept of "guilt beyond a reasonable doubt" inevitably speaks. The concept signifies not any mathematical measure of the precise degree of certitude we require of juries in criminal cases, but a subtle compromise between the knowledge, on the one hand, that we cannot realistically insist on acquittal whenever guilt is less than absolutely certain, and the realization, on the other hand, that the cost of spelling that out explicitly and with calculated precision in the trial itself would be too high (1971: p. 1375).

I don't think that these points should move the defenders of Punish.

Those who defend Punish have two things that they can say in response. First, they might remind us that their argument rests on an assumption that Tribe does not challenge, which is that Punish maximizes expected value. He might think that even if Punish did maximize expected value, we will still see that we ought to follow the norm that he identifies. Why, we might ask, should we do that? If the matrix reflects the value and disvalue that we attach to things like punishing the guilty and punishing the innocent, why should we worry about a *further* ethical consideration having to do with reasonable doubt? We know in advance that conforming to some further norm would sometimes compel us to choose an option we know doesn't serve our values as well punishing someone using statistical evidence would.

We know that the alternative that Tribe defends is a system that allows us to convict using forms of evidence that are less reliable than statistical evidence when it comes to identifying the guilty defendants as guilty and screening out the innocent defendants (e.g., the testimony of witnesses).[6] Because the system that uses of statistical evidence is more reliable in sorting the guilty and innocent into the right category than one that uses a conflicting reasonable doubt standard, it might seem irrational, given our values, to stick to this reasonable doubt standard.

This first response isn't terribly concessive. It simply points out that *if* there is a clash here of the kind that Tribe assumes, there is some reason to think that Tribe's norm is spurious. We know conforming to it wouldn't serve our values as well as using statistical evidence to secure a conviction in violation of the putative norm. There is a second more concessive response to consider. Tribe assumes that there *is* a clash here between a norm that says that we cannot convict if there is reasonable doubt and a norm that says that we should punish if doing so maximizes expected value. Some proponents of Punish might say that this betrays a very strange understanding of

---

[6] Remember that we are operating under the assumption that punishments are sometimes justifiably imposed, which means that juries sometimes have all the evidence they could need to vote to find a defendant guilty and for the criminal justice system to punish accordingly. The evidence that would justify such convictions is typically taken to involve things like testimony from reliable witnesses. In criticizing Punish, most would concede that this evidence is less reliable than the kind of statistical evidence we're discussing and still think that this evidence is adequate. Compare the situation to the lottery situation where writers like Harman (1968) think that we'd be justified in believing things reported in the paper but not justified in believing a ticket to be a losing ticket even though we recognize that the probability that the paper errs is greater than the probability that the ticket won.

reasonable doubt. If we see punishing the innocent as a terrible thing and see failing to punish the guilty as a significantly less bad thing, the probability required for Punish to maximize expected value will be quite high. If it is high enough, shouldn't the statistical evidence be sufficient to justify belief in the guilt of the defendant? If belief in the guilt of the defendant is justified, would that not satisfy any reasonable formulation of the reasonable doubt norm? When judges are asked to translate this talk of reasonable doubt into probabilistic terms, many judges have thought that the relevant probability could be less than .95.[7] If you look at our decision-matrix and consider the probability at which Punish maximizes expected value, we cross that threshold with room to spare.

Tribe has to overcome two challenges to rebut the case for Punish. He needs to show that on the proper understanding of reasonable doubt, the statistical evidence doesn't eliminate it. He then needs to explain why we should prefer a system that conforms to this norm to one that relaxes it and allows us to convict a defendant using statistical evidence when doing so maximizes expected value and thus appears to best serve the values we all share.

3.2 The criminal class

Colyvan, Regan, and Ferson object to the use of purely statistical evidence to secure a conviction because in such a system, 'it *would* appear to be a crime to belong to a reference class' (2001: p. 172). They offer us this example to make their worry vivid:

> …let us suppose that 99 per cent of people from a certain reference class cheat on their taxes. Does this mean that we are justified in charging and sentencing someone in this class with tax evasion, without further evidence? No, of course not; we require more evidence than simply their membership in the reference class in question. It is important to note that we require further evidence *not* because we wish to raise the probability from 0.99 to something higher …Rather, we require further evidence because the reference-class evidence is not specific to the individual in question (2001: p. 172).

Something in this passage seems right. It shouldn't be a crime to belong to a reference class. There is also something very rhetorically effective about framing the issue this way. Having said that, I don't know if this criticism is quite fair to those who defend Punish.

The philosophers who defend Punish don't think that it should be a crime to belong to a reference class. They aren't suggesting that we alter the criminal statutes so that laws that made it a criminal offense to hide income make it a criminal offense to belong to a class composed almost entirely of people who hide income. The philosophers who defend punish think that there is a fallible but adequate test that we can use to decide whether to treat someone as guilty or not and they think that we should use statistical evidence in the test, not in the formulation of the law.

Colyvan, Regan, and Ferson might respond by saying that this distinction between the formulation of the law and a test for its violation doesn't come to much because

---

[7] As noted by Thomson (1986) who cites the research of Simon (1969).

the practical upshot of using a statistical test is the same. That's true, but then it's not clear that this objection has much by way of dialectical significance. I don't think the objection has much dialectical significance because it is hard to imagine an alternative practice to the one that uses statistical evidence that gives them what they desire—a basis for conviction that is specific to the individual in question in a way that statistical evidence is not.

Most of the people who reject Punish accept a view on which juries are justified in convicting defendants on the basis of fallible evidence even when that evidence is misleading. Wrongful convictions are unfortunate, but on this way of thinking, jurors needn't have failed to meet their obligations in convicting an innocent person on the basis of such evidence, provided that the evidential support is adequate. Naturally, the notion of adequacy might be difficult to spell out, but consider Smith's (2016) suggestion that evidence is adequate if it provides *normic* support (i.e., in normal worlds or circumstances, this evidence wouldn't support a false belief). Would it be fair for someone who believes that jurors are justified in convicting the guilty and the innocent when they use evidence that provides normic support in the form of, say, eyewitness testimony or photographic evidence? I think not.[8] This normic support proposal faces an objection nearly indistinguishable from Colyvan, Regan, and Ferson's objection, which is that the view would 'make it a crime' to be the kind of person who could be convicted by juries who used evidence that provided normic support to reach their verdict. But, surely, we might say, it cannot be a crime to be someone who is convicted on the basis of strong but misleading evidence! It would seem that anyone who believes that there can be fallible evidence that justifies juries in convicting an innocent person faces a version of the problem that Colyvan, Regan, and Ferson point to. It would seem that just about everyone believes that such evidence can justify such convictions.[9] If this is a problem for everyone, it really is a problem for nobody. Nobody (with the exception of the objectivists who think that it is always wrong to convict the innocent whatever evidence you have) thinks that the mere fact that a defendant is innocent is itself sufficient to determine that the conviction wasn't justified.

---

[8] Smith (2016, forthcoming) defends this fallibilist view of justification. He is critical of Punish, but doesn't endorse the line of criticism discussed here. Smith's view is designed to vindicate the intuition that statistical evidence of the kind we have in Prisoners is not sufficient to justify belief by virtue of the fact that there are normal worlds in which we have this statistical evidence but end up forming a false belief. It is also designed to vindicate the intuition that evidence that provides normic support can justify belief even if the risk of such evidence leading us astray is greater than some statistical evidence that, on his view, wouldn't justify belief. He embraces the conclusion that, say, experiential evidence can justify a belief even when it's more likely that beliefs based on this evidence are mistaken than it would be if they were based on some statistical evidence.

[9] This is in line with Schauer's (2003: p. 86) suggestion that the problem that we're dealing with isn't about statistical evidence, per se, but about the use of nonuniversal indicators. Normic support would seem to be just one more nonuniversal indicator and while there might be reasons to think that some of these indicators are better suited to the task of a criminal trial, I don't think that the point about reference class is helpful since we can create reference classes by reference to the classes picked out by the relevant indicators (e.g., the class of individuals who we could believe to cheat on their taxes using evidence that provides normic support).

3.3 Thomson's guarantee

In her discussion of the puzzle of the proper role of statistical evidence in the law, Thomson (1986) notes something that seems to be operative in the passage just quoted, which is that the purely statistical evidence seems to be not properly about or connected to the defendant. What we desire, she suggests, is evidence that is causally connected to the defendant in question. On her view, just conviction requires *individualized evidence*, evidence that is causally connected to the defendant and the defendant's deeds in a way that statistical evidence wouldn't be.

Thomson's suggestions point us in the right general direction, but the rationale she offers for her view is problematic. Why should the law care about the difference between statistical evidence and individualized evidence? Thomson says two things about individualized evidence to convince us that we shouldn't convict without it. First, she says that such evidence can provide a guarantee that statistical evidence cannot. Second, she says that such evidence can be the basis for knowledge and that statistical evidence cannot.[10] These points might be correct, but do these epistemological points support her objection to Punish?

Enoch, Fisher, and Spectre think that the law shouldn't concern itself with these epistemological matters:

> …why should the law of evidence care about knowledge or about epistemology more generally? It should care, undoubtedly, about truth, accuracy, and the avoidance of error. But why is it important that courts base their findings on knowledge? Insisting that the law should, after all, accord significant weight to knowledge or to epistemology in general amounts to a willingness to pay a price in accuracy. Indeed, excluding statistical evidence amounts to excluding what may be genuinely probative evidence. And this means that the legal value of knowledge—if it has legal value and if that value is what grounds the differential treatment of statistical and individual evidence—sometimes outweighs the value of accuracy; that, in other words, in order to make sure that courts base their ruling on knowledge, we are willing to tolerate more mistakes than we otherwise would have to and, in fact, a higher probability of mistake on this or that specific case. This just seems utterly implausible (2012).

Thomson has a response, but I fear that it only serves to highlight the difficulty we face in trying to resist the argument for Punish:

> Our society takes the view that, in a criminal case, the loss to society if the defendant suffers the penalty for a crime he did not commit is very much greater than the loss to society if the defendant does not suffer the penalty for a crime he did commit …This point might be re-expressed as follows: our society takes the view that in a criminal case, the society's potential mistake-loss is very much greater than the society's potential omission-loss. It would be no wonder, then, if our law

---

[10] It is unclear what she means by a guarantee in this context. She might mean that it is a kind of connection to the truth that we would get from a causal connection. The context suggests that she thinks that such a connection is needed for knowledge and is missing in cases where we form beliefs based on statistical evidence alone.

imposed a heavy standard of proof on the jury in a criminal case; and according to the friend of individualized evidence, that means the jury must be very sure of having a guarantee before imposing liability for a crime (1986: p. 215).

In my view, this is a strange argument. We'd expect the proponents of Punish accept everything up to the last sentence. Their argument is based on the same observations about value. These observations about value are observations they offer in support of Punish. They'd say that the kind of guarantee that Thomson is after is either something the law shouldn't care about or something that the statistical evidence provides. It looks as if Thomson is saying that we shouldn't Punish because we recognize that the society's potential mistake-loss is greater than the society's potential omission-loss. The proponents of Punish would say that we *should* Punish because of these points about value, provided that the probability of guilt is sufficiently high. When the probability of guilt is sufficiently high, the proponents of Punish think that Punish best serves the very values that Thomson appeals to in arguing against Punish. Because she identifies no clear rationale for thinking that we should choose options that have less expected value than Punish, it isn't clear that there's any rationale here for insisting that it is never just to punish using statistical evidence.

There's something interesting about Thomson's response to the puzzle that I want to flag. She thinks that there is something epistemically bad about believing the defendant to be guilty and that this epistemological fact explains why Punish is mistaken. To meet the Enoch, Fisher, Spectre challenge, we have to show that the law should care about the epistemological matters that matter to Thomson. We also have to show that Thomson's right about the epistemology. In the next section, we'll see that there's a *prima facie* plausible argument for thinking that there isn't anything epistemically wrong with believing the defendant to be guilty in Prisoners.

4 An argument for believe

Just as there is disagreement about whether to punish using statistical evidence, there is disagreement about whether such evidence justifies a full belief in guilt. There is a *prima facie* plausible argument for Believe that parallels the argument for Punish. Believers want to acquire true beliefs and avoid false ones. Our desires and aversions reflect the epistemic value we assign to accurate and inaccurate belief. We can represent these values using a belief-matrix that looks similar to the decision-matrix from above:

Veritist Belief-Matrix

	Guilty	Innocent
Believe	+1	−10
Don't Believe	0	0

If we accept a veritist view on which truth and falsity are the fundamental epistemic goods, we might say that it is a good thing to correctly believe the defendant to be

guilty and a bad thing to mistakenly believe the defendant to be guilty.[11] On this view, we could say that there is nothing good or bad that comes from having no belief. If this matrix captures our epistemic values and we can assign probabilities to the states, we can determine what it takes to maximize expected epistemic value.

With this much in place, we can now offer an argument for Believe:

An Argument for Believe

P1'. Rationality requires us to maximize expected epistemic value.[12]

P2'. Believe is the option that maximizes expected value.[13]

P3'. So, rationality requires us to believe.

P4'. It would be right to believe what rationality requires.

P5'. So, it would be right to believe.

If the argument for Believe is sound, this undermines the epistemological objections to Punish.[14] Readers should now see why I'm concerned with the extant case against Punish. It isn't clear which premise, if any, we should reject in the argument for Punish if we reject Believe. It isn't clear what right we have to assume that Believe is mistaken given that Believe appears to best serves our epistemic values.

5 Moral considerations

We now have our arguments for Punish and Believe on the table. If the reader's moral sensibilities are anything like mine, they'll be troubled by the suggestion that we can punish using statistical evidence. They might try to identify some moral principle that we would violate if we were to punish in Prisoners. I shall discuss three.

---

[11] For defenses of this view of epistemic value, see Lynch (2004). Some sophisticated veritists (e.g., Sylvan forthcoming) think that the fundamental epistemic goods are truth and falsity and also think that knowledge is better from the epistemic point of view than mere true belief. We will look at these views later. For a nuanced discussion that defends the view that knowledge is not better from the epistemic point of view than true belief, see Pritchard (forthcoming).

[12] This is assumed by many epistemic consequentialists. See Dorst (forthcoming) and Easwaran (2016) for decision-theoretic justifications of Lockean views of rational belief. For defenses a quasi-consequentialist justification approach to epistemic norms, see also Joyce (1999) and Pettigrew (2016). For arguments that these quasi-consequentialist approaches are implausible because they would justify problematic tradeoffs, see Berker (2013), Carr (2015), Firth (1981), Jenkins (2007), and Littlejohn (2012, 2015). Not everyone is impressed with this line of objection. See Ahlstrom-Vij and Dunn (2014) and Talbot (2014). While I think that the quasi-consequentialist approach is problematic, my response to these arguments will focus on the veritist value theory, not the quasi-consequentialist norms.

[13] This is plausible if veritists like Joyce (1999) are right that accuracy and inaccuracy are the fundamental epistemic goods.

[14] Philosophers attracted to the Enkratic Requirement (i.e., the requirement that states that we should see to it that we φ when we believe that we ought to φ) should notice that they can use the Enkratic Requirement to connect Believe to Punish more directly. If the Enkratic Requirement is correct, Punish couldn't be mistaken if Believe is correct. For defenses of the Enkratic Requirement, see Gibbons (2013), Littlejohn (2012, forthcoming), and Titelbaum (2015).

I accept the first, but know that most readers do not share my objectivist instincts. Some readers might be attracted to the second, but I think it is problematic. My case against Punish ultimately rests on the third. We shall see that that this third argument is sound iff we can establish that Believe and Punish are connected and that Believe is mistaken.

5.1 Never the innocent

Because I have objectivist instincts, I think that we should conform to this norm:

> Never the Innocent: It isn't permissible to punish people for crimes they didn't commit.

This norm is an objectivist norm in the sense that it has an objective application condition (i.e., one that doesn't supervene upon our subjective states, individually or collectively as jurors). This norm follows from a standard reading of a knowledge norm that states that we should not convict a defendant unless we know them to be guilty.[15]

If Never the Innocent is a genuine norm, the primary argument for Punish must be unsound.[16] To see why, think think about what happens if we selected the one innocent prisoner. If our defendant had been innocent we would still maximize expected value by punishing the defendant. Thus, the argument for Punish is, *inter alia*, an argument for violating Never the Innocent. Since Never the Innocent is a norm we shouldn't violate, we should reject Punish.

5.2 Capped convictions

Unfortunately, most readers probably don't find the previous argument convincing because it rests on the objectivist assumption that the defendant's innocence is sufficient to establish that it would be wrong to punish. They might think that there is something in the neighborhood of this argument against Punish, however, that works. Notice that if the argument for Punish works in one case, it would work in each case. If it worked in each case, we'd have a case for thinking that we should punish 100 prisoners. This, however, might seem outrageous. It might seem outrageous to punish 100 people for 99 crimes.

This argument assumes this intuitively plausible principle:

> Capped Convictions: It is never permissible to knowingly punish N people for N-1 offenses.[17]

---

[15] See Blome-Tillman (MS) for a defense of this objectivist norm.

[16] For defenses of this kind of objectivist principle, see Littlejohn (2012).

[17] Peter Dennis first raised this objection during a presentation of Katie Steele (MS). I initially thought the objection was decisive but was brought around by discussions with Julien Dutant, Martin Smith, and Katie Steele.

While the argument that appeals to this principle might initially seem promising, I doubt that this argument could succeed if we are right to reject objectivist norms like Never the Innocent.[18],[19]

To see why, remember that the people who reject Punish think that we're sometimes justified in punishing defendants. They would presumably accept some principle along these lines:

Any Sure Offender: It is permissible to punish a defendant when we justifiably believe the defendant to be guilty.[20]

This case shows that Any Sure Offender should lead us to reject Capped Convictions:

Prisoners II

100 prisoners have been convicted. In each case, the juries relied on adequate evidence for the belief that the defendant was guilty. After the convictions were carried out, a perfectly reliable observer tells you that precisely 1 of these people had been framed. Alas, your informant dies before he can identify the person.[21]

While the objectivists who accept Never the Innocent could say that there is precisely one person in Prisoners 2 who should be freed (and one person who never should have been convicted), we're operating under the assumption that this objectivist view is mistaken. If we accept Any Sure Offender, we'll have to reject Capped Convictions. The testimony might put us in a position to know that there are only 99 guilty people in the cells but this doesn't prevent us from justifiably believing in each of the 100 cases that the particular defendant we're considering is guilty.

5.3 Reasonable conviction

To my mind, the strongest (non-objectivist) objection to Punish focuses on the relationship between belief and punishment.

---

[18] Much in the way that consistency norms follow from truth norms, Never the Innocent vindicates Capped Convictions. Thus, *if* you thought that Never the Innocent was a genuine norm, you would have to accept Capped Convictions.

[19] An anonymous referee noted that there is a variant argument for Punish that Never the Innocent does not block. The argument starts from the observation that it is very probable that you should believe the defendant to be guilty. With this assumption and the further assumption that it is very probable that you should do something if it is very probable that you should believe you should do it, we get the result that it is very probable that you ought to punish the prisoner. In the later sections, I shall argue that it is *not* very probable that you should believe the defendant to be guilty on the grounds that probable truth is sufficient for neither a justification to believe nor for making it probable that you should believe a proposition. That is because, as we'll see, it might be very probable that p and yet you might know that nobody could know whether p. When you know this, I think that it is settled that you should not believe p and that it is *not* probable that you should believe p.

[20] We should assume that this principle uses a non-factive notion of justification so that it could recommend convicting someone who is innocent if, say, we had the right kind of evidence for believing the defendant to be guilty.

[21] The adequacy of evidence has to be understood in such a way that it is possible to have adequate evidence to believe falsely that the defendant is guilty.

Remember our first matrix, the Veritist Decision-Matrix. To determine whether Punish maximizes expected value, we needed to know the probability of guilt and the values of <Punish, Innocent> and <Punish, Guilty>. For some assignments of value, the probability of guilt has to be very high for Punish to maximize expected value. If it's at all plausible to think that it's rational to believe a proposition when its probability on the evidence is sufficiently high, it might seem that rational beliefs about guilt and rational decisions about whether to punish will go hand in hand. The crucial point to notice, however, is that this correlation doesn't hold for many possible value assignments to <Punish, Innocent> and <Punish, Guilty>.

Blackstone said that it is better that ten guilty persons walk free than one innocent person suffers. Franklin said that it would be better that one hundred should walk free than one innocent person be sent to prison. With numbers like these, there's some correlation between the punishment that maximizes expected value and high probability of guilt. Voltaire, however, thought that it was much more prudent to let two guilty persons walk free than to let one person suffer. With numbers like this, we lose the connection between high probability and guilt. If readers prefer the views of Joseph Stalin or Dwight Schrute, they'd think that it would be better that hundreds if not millions of innocent people be jailed than it would be to let one guilty person walk free. On *these* assignments of value, the justification of punishment wouldn't require the probability of guilt to be particularly high.

This points to a general concern. Let's say that the probability of the defendant's guilt is at the *practical threshold* iff the expected value of Punish is equal to the expected value of Don't Punish. The probability of the defendant's guilt is at the *epistemic threshold* iff the expected value of Believe is equal to that of Don't Believe. In setting out the case for Punish, it seems that none of the operative assumptions would ensure that the practical threshold is at least as high as the epistemic threshold. If we accept the epistemological assumptions operative in the argument for Believe, we would be committed to the view that it is reasonable to believe (outright) that the defendant is guilty if the probability of guilt is at least as high as the epistemic threshold and that it would be unreasonable otherwise. Thus, it's possible that the best argument for Punish is an argument that we ought to Punish even if we know that it's unreasonable to believe that the defendant is guilty.

To block this, someone could argue that if we plug the right values into the Veritist Decision-Matrix, the practical threshold is very high, high enough to match or surpass the epistemic threshold. I have two worries about this response. First, it wouldn't be in keeping with the view to insist that, say, Franklin's numbers were in the right ballpark on the grounds that it raised the practical threshold to a level that would satisfy our epistemic scruples (i.e., a point at which the probability of guilt is sufficiently high to convince us that outright belief in the defendant's guilt would be reasonable). Our assignment of values to the Veritist Decision-Matrix wasn't initially driven by epistemological concerns (e.g., a concern that the practical threshold is too low to ensure that jurors who rightly convict a defendant could reasonably believe the defendant to be guilty). We were supposed to justify a decision to convict without appeal to epistemological considerations by focusing on the values we wanted our criminal trials to promote. If we want the decision-matrix to reflect these values, we might need to appeal to assumptions about reasonable belief that haven't yet played any explicit role

in the discussion. In turn, we might think there was something wrong with the way we initially framed the decision problem. Second, we might reject the very idea that the epistemic threshold represents something significant when it comes to rational full belief. *If* we accept the argument for Believe, we might think that this threshold represents the weakest level of support at which it is rational to believe and we might then wonder how this point relates to the practical threshold. If, however, we reject the Lockean view of rational belief, we might worry that it could be irrational or unreasonable to believe a defendant to be guilty even if the probability of guilt exceeds both the practical and epistemic threshold.

Can the proponents of Punish give us any reason to think that they can vindicate this norm?

> Reasonable Conviction: It is not permissible to punish a defendant if it isn't reasonable to believe the defendant to be guilty.

I don't think that they can. If the potential loss to the defendant is small enough or the importance of punishing the guilty is great enough, we should expect that there could be situations where it is irrational to believe outright that the defendant is guilty even though Punish would maximize expected value (as characterized by the Veritist Decision-Matrix). If the proponents of Punish had said up front that they were defending a view on which we can be required to punish a defendant even when we know it would be unreasonable to believe outright that the defendant to be guilty, their argument wouldn't have had the persuasive force it initially appeared to have. I think that if a member of the jury, say, told us that they were going to vote to convict in spite of the fact that they thought they had to suspend judgment on whether the defendant was guilty, we wouldn't be impressed by their use of a decision-matrix to show that Punish maximized expected value.[22]

If the proponents of Punish try to show that Punish maximizes expected value only when the probability of guilt is very high, we might reasonably worry that their axiology was being driven by epistemological considerations. If that's how things should work, the law should take an interest in the epistemological considerations that Enoch, Fisher, and Spectre thought shouldn't interest the law. It also looks like the initial Veritist Decision-Matrix would be too crude to take account of the values we want to plug in (e.g., the matrix doesn't distinguish between the case where we convict the guilty while irrationally believing them to be guilty and the case where we convict while reasonably believing the defendant to be guilty). Once we let the epistemic considerations help determine which values should go into our matrix, we should insist on a matrix that divides up the states and options differently.

Since the primary argument for Punish *is* an argument that succeeds only if it would show that it would be right to punish someone who we know we could not reasonably

---

[22] An anonymous referee asked why this is. Why wouldn't it be enough that the agent was just very confident that the defendant was guilty? I shall explain why in the next section. The important point is that deciding to punish isn't like deciding to take an umbrella. For reasons that I shall discuss below, reactive attitudes and affective responses make a difference here and force us to see a difference between betting on the guilt or innocence of a defendant and punishing a guilty defendant.

believe to be guilty, I think we know there is something wrong with the rationale offered in support of Punish.

Why should we think this? I don't think it's because we're convinced that the right values are closer to something like the values Franklin suggested. We don't come to this issue thinking that the right values to plug into the cell *must* ensure that Punish only maximizes expected value when the practical and epistemic threshold match. I think we are troubled by the rationale offered for Punish because on a proper understanding of the values at play, the values we assign to our cells are sensitive to epistemological considerations. The Veritist Decision-Matrix doesn't reflect this, but this is a reason to think that there's something wrong with the way the issue has been framed. The debate has been framed as one in which one side thinks that we ought to punish because we ought to maximize expected value and the other side thinks that we shouldn't punish even though they seem to grant that Punish does maximize expected value. It was a mistake to do so. I shall have more to say about this and alternative decision matrices in Sect. 7.

My main argument against Punish can be stated as follows:

The Argument against Punish

A1. It would be wrong to punish the defendant in Prisoners if we could not rationally believe the defendant to be guilty.

A2. Given the grounds in Prisoners, we could not rationally believe the defendant to be guilty.

Ac. Thus, it would be wrong to punish the defendant in Prisoners.

The first premise, (A1), is just Reasonable Conviction. The second premise is an epistemological claim that we'll discuss below.

Some would likely say that if Reasonable Conviction supports an argument against Punish, it wouldn't be a genuine principle. If the principle supports the case for Punish, wouldn't it tell us that we shouldn't maximize expected value? Don't we have good reason to think that we ought to maximize expected value and thus violate putative principles that would tell us to do otherwise? If so, it clearly calls for justification.

This is a fair request. My defense of Reasonable Conviction begins with a reminder that punishment is an act that differs in an important way from acts like betting on football matches or taking umbrellas.[23] This is because punishment is supposed to be a way of holding someone accountable and it involves a backwards-looking element that other actions often lack. Thus, the act in question (e.g., imposing a prison sentence) has to be guided by certain kinds of consideration to be a punishment. If Agnes' reason for

---

[23] Adler (2002) and Buchak (2013) have argued that the grounds needed for blame differ from the grounds needed for betting. A high degree of confidence might make it rational to bet on Chelsea (depending upon the odds) even if this high degree of confidence doesn't amount to full belief. (Moreover, a low degree of confidence could also make it rational to bet if the odds are right.) A degree of confidence that doesn't amount to full belief does not rationalize reactive attitudes like resentment. As Adler memorably puts the point, 'Mild resentment is never resentment caused by what one judges to be a serious offense directed toward oneself tempered by one's degree of uncertainty in that judgment' (2002: p. 217). None of this is surprising if we see knowledge as required for factive emotion. See Dietz (forthcoming) and Gordon (1987).

harming Jack couldn't be anything to do with what Jack has done, Agnes' act wouldn't properly be described as punishment. Thus, we need to foreground something that has been left in the background too long. If we have a system of rules that governs decisions to punish or to refrain from punishing, it would seem that the rules should require that the decision to impose the harms associated with punishment be made only when the punishment can properly express blame or at least treat the defendant as accountable for some specific deed. It would not be proper to blame unless the relevant parties could properly believe that the defendant did something blameworthy. This is why we shouldn't adopt rules that tell us to make a decision that would be expected to harm a defendant on the basis of considerations of expected value if the agents imposing those harms couldn't also believe something about the defendant's guilt. Recall at the beginning that I said that any just system of punishment would have to have a backwards-looking element. It would have to have something that would ensure that there was a prohibition against imposing the harms by convicting for any reason that didn't include the defendant's guilt.[24]

This is why I think that it's plausible that a norm like Reasonable Conviction should be operative in criminal trials. There is no obstacle to justifying acts like taking an umbrella or making a bet by appeal to considerations of expected value because the justification of these acts doesn't turn on whether the agent has any specific full beliefs about the situation. An agent can justifiably take an umbrella because they believe outright that it is raining, but they could also take an umbrella because they fear that it might and have a strong aversion to getting wet on the way to work. When we're dealing with acts that have an expressive dimension, the justification of the act turns, in part, upon the justification of the relevant accompanying attitudes. Without the relevant attitudes, no act could serve that expressive function. Without a belief in the defendant's guilt, we could not act in a way that would express blame. Our actions wouldn't be ways of holding the defendant to account. We would (hopefully) object to the suggestion that the jury would rightly conclude that a defendant should be forced to suffer a harm if we didn't think that the defendant was guilty of the act that was supposed to merit the punishment.

6 Believe?

If Reasonable Conviction is a genuine norm, my argument against Punish will succeed iff Believe is mistaken. If we can reasonably believe that the defendant in Prisoners is guilty, there is no principled objection to conviction. If, however, it's not reasonable to believe the defendant to be guilty, Reasonable Conviction tells us that it would

---

[24] As an anonymous referee pointed out, a focus on blame might be too specific for my purposes. We can offer an account that parallels Buchak's (2013) even if punishment does not involve blame. All that matters for our purposes is that punishment is like blame in that it involves a backwards-looking element. The crucial idea is that this backwards-looking element is responsible for the crucial dependence of the response (i.e., blame, punishment) upon certain specific beliefs about the situation (e.g., the agent's deeds, the agent's attitudes at the time of action). If, say, punishment involves at least some backwards-looking element, it isn't surprising that the justification of it might depend, in part, upon the justification of certain specific motivating beliefs.

be wrong to punish even if Punish turns out to be the best option on the Veritist Decision-Matrix.

In this section, we'll look at three approaches to rational belief to try to determine whether Believe is correct.

6.1 Veritism and the Lockean view

The argument for Believe should appeal to Lockeans about rational belief (i.e., those who think that a full belief is rational iff it is rational for the thinker to have a sufficiently high degree of confidence). The argument shows that we maximize expected epistemic value (characterized in terms of accuracy and inaccuracy) iff we believe the defendant to be guilty on the basis of the statistical evidence.

There are serious problems with this Lockean conception of rational belief and the quasi-consequentialist arguments that support it. The Lockean view doesn't vindicate the intuitions that many of us have about beliefs about lotteries and beliefs based on testimony. While I don't think that it's rational to believe (fully) that a ticket in some lottery lost when that belief is based on statistical evidence, it can be rational to believe what a reliable newspaper reports even when it's more likely that the paper errs in this entry than it is that the ticket wins.[25]

I wouldn't expect these considerations to trouble the Lockean, but perhaps the second problem will cause more concern. There are certain 'bad' propositions that we know that we cannot know where these propositions are at least as probable on the evidence as propositions that the Lockean takes to be rational to believe. Suppose we know, for example, that we cannot know whether lottery propositions are true. It doesn't seem rational to believe this: I don't know if my ticket lost, but it did. There is a kind of incoherence or clash present here in which the thinker seems to be both putting it forward that they were aware of something while affirming that they are not. In spite of the apparent irrationality of believing this, the proposition should be as probable on my evidence as the proposition that my ticket lost. Thus, on the Lockean view, it seems that these propositions should be equally rational to believe. If they want to vindicate the intuition that it's not rational to believe the Moorean absurdity, they have to raise the probability threshold necessary for rational belief to absurdly high levels. If they use quasi-consequentialist reasoning to test epistemic norms, they would have to modify their value theory in such a way that they'd show an aversion to risk that is pathological. It's worth pointing out here that if it's irrational to believe Moorean absurdities (e.g., that God hates my atheism, that my ticket will lose but I don't know if it will), it's irrational to believe that which we know we cannot know. If rationality doesn't permit us to believe both conjuncts, it won't allow you to believe one conjunct (e.g., my ticket will lose) by trying by hook or by crook to suppress the knowledge that this isn't something you're in a position to know.

It also seems that the Lockean view will never vindicate our intuitions about cases like Prisoners and Prisoners II. In Prisoners II, it's rational to believe that some randomly selected prisoner was guilty. In Prisoners, it is not. From the Lockean per-

---

[25] I think this observation is first due to Harman (1968).

spective, the two cases should be on par since the truth of both beliefs is highly probable on the evidence. It seems that our intuitions draw distinctions between cases that don't pattern with differences in the probability of the relevant beliefs.

In short, the Lockean view tells us a plausible story about what rational belief would be like if we accepted a veritist theory of epistemic value and thought that rational belief, like rational action, is rational because of how it promotes that value, but I fear that the view's verdicts are implausible in a number of cases. While it is clear that there is a Lockean argument for Believe, I don't think we should accept the Lockean approach.

6.2 Defeatism

While some veritists are attracted to the Lockean view, some would argue that the Lockean account runs into trouble because it doesn't take account of potential defeaters that could make it irrational to believe highly probable propositions. To undermine the case for Believe, some readers might be tempted by some such defeatist response. A defeatist would say that there is some sort of epistemic principle that provides us with a defeater that defeats the justification provided by the thinker's evidence in cases like Punish. Concerning Prisoners, the defeatist could say that while we can rationally be very confident that the defendant is guilty we cannot justifiably or reasonably believe that a defendant is guilty because this belief would violate this principle:

> Avoid Falsity Principle: For any set L of competing statements, if (i) a person S has good reason to believe each member of L is true and (ii) either S has good reason to believe at least one member of L is false or S is justified in suspending judgment about whether at least one member of L is false, then S is not justified in believing any of the competing individual members of L (Ryan 1996: p. 130).

As in standard lottery cases, Prisoners is a case in which we know that if we believe every claim supported by the strong statistical evidence we will believe one falsehood. Thus, the Avoid Falsity Principle tells us that we cannot rationally believe any of the defendants to be guilty. This would show that Believe is mistaken. In turn, I could appeal to Reasonable Conviction to show that Punish is mistaken.

While I think that this defeatist verdict concerning Believe is correct and that we ought to reject the Lockean view that says that high probability of truth is sufficient for justified belief, I think the defeatist response suffers from two problems. The first is that it seems to deliver the wrong verdict in preface-type cases. By my lights, it delivers the right verdict in Prisoners, but the wrong verdict in Prisoners II. It is like the Lockean view in treating these cases equally but what we want is a view that treats them differently.[26] In Prisoners II, it seems plausible that the thinker is justified in believing each prisoner to be guilty even though the Avoid Falsity Principle predicts that this justification should be defeated. Much in the way that it seems to be an

---

[26] Ryan (1996) offers a defeatist treatment of preface-type cases. For defenses of rational inconsistency, see Christensen (2004) and Easwaran and Fitelson (2015), My defense of rational inconsistency appeals to Bird's (2007) knowledge-first theory of rational belief. See below.

overreaction to the discovery that one prisoner was framed to release all the prisoners, it seems to be an overreaction to the discovery of error to suspend judgment on a large set of propositions each of which is a good candidate for knowledge or for knowledge-level justification.

The second problem with the defeatist line is that the principle is likely to strike people as *ad hoc*. There is something compelling about what Horowitz says about rationality:

> …a rational agent should be doing well by her own lights, in a particular way: roughly speaking, she should follow the epistemic rule that she rationally takes to be most truth-conducive. It would be irrational, the thought goes, to regard some epistemic rule as more truth-conducive than one's own, but not adopt it (2014: p. 43).

If the values in the Veritist Belief-Matrix are *our* values, doesn't it seem that these values are best served by a set of rules that tolerates inconsistency in large sets of beliefs? The defeatist doesn't go far enough because the defeatist doesn't question the veritist assumptions that underwrite the arguments for the Lockean view. We know that the rules that tolerate belief based on statistical evidence, if followed, would do a better job leading us to form true beliefs and avoid false ones than alternative sets of rules that includes rules like the Avoid Falsity Principle. It thus seems that it would be irrational for someone with the veritist's values to adhere to that principle. The parallel with the practical case is instructive here. (This echoes my concern about Thomson's rationale for thinking that just punishment requires a guarantee from individualized evidence.) I don't think that many people would think that a rational actor would accept some analog of the Avoid Falsity Principle and use that to guide their choices (e.g., by refusing to convict any defendant in cases like Prisoners II or refusing to make a dutch book against someone because they knew that this sure-win strategy would require losing a bet).

6.3 Gnosticism

We've looked at two views of rational belief. They offer different verdicts concerning Believe, but I think that they're problematic because they don't distinguish Prisoners from Prisoners II. Luckily, there is an alternative.

In discussing the Lockean and defeatist views, we assumed a veritist value theory on which accuracy and inaccuracy are the fundamental values. With this evaluative framework in place, we seem forced to choose between an approach that identifies justification with high probability or a framework that introduces defeaters that seem *ad hoc* and would appear to defeat too much justification. We should consider an alternative approach to epistemic value.

Instead of treating accuracy and inaccuracy as the fundamental epistemic goods and evils, the *gnostic* treats knowledge and failed attempts at knowing as the fundamental

epistemic values.[27] Because there are beliefs that are highly probable on the evidence that we know aren't things that we can know (e.g., lottery propositions, Moorean absurdities), the gnostic isn't tempted to say that these beliefs are justified or rationally held. If we know *apriori* that such beliefs are epistemically disvaluable, we don't need to introduce defeaters to explain why we shouldn't hold them.

We know why the gnostics wouldn't think of high probability as sufficient for rationality, but we don't know yet what they take rational belief to be. There are at least two promising approaches to consider. Nothing in this paper would turn on which approach we pursued. One approach would be similar to the veritist Lockean view in that it would say that a rational belief is rational because it maximizes expected epistemic value. It differs from the veritist-motivated Lockean view because it would tell us that the rational status of a belief turns on the probability on the thinker's evidence that the belief would be *knowledge*, not (just) on the probability on the thinker's evidence that the belief would be *accurate*. Suppose we thought that knowledge required some modal condition (e.g., safety, sensitivity) or some causal condition that connects that which makes the belief true to the belief itself. In lottery-type cases, we would know that high probability of accuracy and high probability of constituting knowledge would come apart because the target belief wouldn't meet the relevant modal or causal condition. In such cases, the gnostic thinks we ought to suspend judgment in spite of the high probability of the target belief being true.

There are alternative approaches to rational belief that the gnostic might consider. Recall Bird's suggestion that rational belief is a kind of counterpat of knowledge:

> Knowledge is epistemically central. Justified belief is a certain kind of approximation to knowledge. It is an approximation that is independent of one's mental states. If one attempts to know but fails for some reason that is located outside one's mental states then one's belief is justified. But if one's failure to know is due to some feature internal to one's other mental states, then the judgment is not justified (Bird 2007: p. 84).

Whenever your beliefs constitute knowledge, they are rationally held. If you believe but don't know, your beliefs will count as rational if the failure to know isn't down to the way that you've exercised your rational capacities. In lottery cases, the failure to know is internal to the thinker's mental states—their belief is supported by the wrong kind of considerations. In preface cases, however, there is no such failure to point to—every belief in the preface-type case might be an approximation to knowledge.

The differences between these two ways of developing a gnostic view of rational belief shouldn't matter for our purposes because they both agree that if we know *apriori* that our belief in p could not constitute knowledge, we know *apriori* that we cannot rationally believe p. Both of these approaches give us a principled basis for rejecting Believe.

Gnosticism fares better than the veritist Lockean view or the defeatist view because the gnostic sees some accurate states of mind as states that realize the fundamental

---

epistemic evil and fail to realize any sort of compensating epistemic good. Concerning the epistemic options in Prisoners, for example, gnostics could offer us this revised matrix: With a matrix like this, the case for Believe isn't just blocked; we have a case

Gnostic Belief-Matrix

	Known Guilt	Unknown Guilt	Innocent
Believe	+1	−10	−10
Don't Believe	0	0	0

against Believe.[28] Given the gnostic account of epistemic value, Don't Believe does better than Believe.

This chart summarizes the verdicts of the three approaches to rational belief considered thus far: The defeatist gets lottery propositions and Prisoners right, but the

	Lottery/Prisoners	Preface/Prisoners 2	Moorean Absurdities
Lockean	Rational	Rational	Rational
Defeatist	Not Rational	Not Rational	?
Gnostic	Not Rational	Rational	Not Rational

tool that it uses to justify these verdicts is too crude. The defeaters introduced prevent us from distinguishing the lottery from the preface or Prisoners from Prisoners II. The Lockean view is too crude, too. It gets the preface and Prisoners II right, but it also doesn't distinguish between these cases and the lottery or Prisoners. The gnostic has no trouble distinguishing between these cases because those in the first column are things we know *apriori* cannot be known and those in the second column are plausible candidates for knowledge. The gnostic approach is the only view we've seen that gets all the cases right. It explains why we should reject Believe and so why we should see Reasonable Conviction as a problem for Punish.

7 The gnostic solution(s)

One reason that the primary argument for Punish and the argument for Believe seem compelling is that these arguments seem to simply combine some observations about the things we value with some norms taken from decision-theory and churn out a result. If these values are our values and we don't want to bet against decision-theory, it seems we are stuck with Believe and Punish. It appears, for example, that Punish is a straightforward case of decision under risk. We know the probabilities and the

---

[28] An anonymous referee asked whether it is plausible to think of true beliefs that fail to constitute knowledge and false beliefs as having the same disvalue. I think that the answer is 'Yes'. Both beliefs fail to do what full beliefs are supposed to do, which is to put us in a position to believe, feel, and do things for a reason that is constituted by a fact that the knowing agent has in mind. See Littlejohn (forthcoming b). All that matters for our purposes, though, is that it is worse to believe and fail to know than it is to suspend.

values of the outcomes. The objectively best and worst outcomes are uncertain, so it seems that we should just stick to the rule that says that we maximize expected value. The problem with appealing to norms like Reasonable Conviction is precisely that it seems to require that we act against this rule.

It is this apparent clash between a norm like Reasonable Conviction and the norms of decision theory that made it difficult to see how Tribe and Thomson rejected Punish. I would now like to suggest that there is a way in which this clash is merely apparent. If we think that the values that the law ought to care about are those that figure in the Veritist Decision-Matrix or Belief-Matrix, the clash would be real and we would have to choose between Reasonable Conviction and maximizing expected value. If, however, the values that the law ought to care about are those that figure in a Gnostic Decision-Matrix, there is no clash:

Gnostic Decision-Matrix

	Known Guilt	Unknown Guilt	Innocent
Punish	+1	−10	−10
Don't Punish	−1	0	0

If we frame our decision problem in this way, the norm that tells us to maximize expected value tells us that we shouldn't punish. Given these values, this norm delivers the same verdict as Reasonable Conviction. Indeed, the decision problem is, from this point of view, not a case of decision under risk. For this to be a case of decision under risk, there has to be uncertainty about which outcome is objectively best. There isn't. We know *apriori* that we're in a case where nothing good could come from Punish and that we are certain to avoid a bad outcome if we choose Don't Punish.

If I'm right in suggesting that the right values are closer to those contained in the Gnostic Decision-Matrix, we see that the problem with the primary argument for Punish is not the application of norms from decision theory, but some assumptions about value that hadn't been questioned in earlier discussions of the problem posed by the use of statistical evidence in criminal trials. Readers will undoubtedly want to know why we should think that the law ought to care about the difference between punishing someone known to be guilty and punishing someone when there is a very high probability of guilt.

My answer to this question draws on an observation from Sect. 5. Recall Reasonable Conviction. While the criminal trial is designed to help identify the guilty and screen out the innocent, it isn't concerned with just the efficient and reliable sorting of people into groups. In one outcome of the process, the defendant is held accountable for some wrong that they've committed. The imposition of punishment is supposed to serve this function and the actions that serve this function are not just the actions that result in a reliable distribution of defendants into the right piles. Thus, the trial process should be regulated to ensure that if the trial is properly resolved with a decision to punish this punishment could fulfill the relevant expressive function. It can do so if the relevant agent's reason for imposing some harm is the fact that the defendant is guilty of the offense. It cannot do that, however, if that fact could not be the agent's reason, say, for

imprisoning someone. An agent cannot blame someone for having done something if they know that they don't know that the defendant did the deed. This is because an agent's reason cannot be something they know they do not know.

The law should assign different values to two possible outcomes of a criminal trial:

(a) Knowingly inflicting harm upon someone who cannot be blamed for wrong-doing;

(b) Knowingly inflicting harm upon someone who can be blamed for wrongdoing.

It is a perversion of the process to impose harm upon someone when we aren't in a position to hold the person accountable for her deeds. Notice that the original Veritist Decision-Matrix lumps these two outcomes together. If I'm right about Believe, the success of their argument for Punish turns on lumping these two things together. (If they were to concede, for example, that it's not reasonable to believe in Prisoners that the defendant is guilty, they would have to concede that it's not reasonable to blame the defendant in this case. If they conceded that and then assigned different values to (a) and (b), they would need to modify their decision-matrix so that it more closely resembled the Gnostic Decision-Matrix.)

Where does knowledge enter the picture? It can enter the picture in two ways. It matters whether the agents responsible for decisions about conviction and punishment are able to make a decision to impose the harms *for the reason that* the defendant committed some offense.[29] If they cannot do this, it isn't reasonable to blame. If they can do this, it is reasonable to blame. On the first way of bringing knowledge into the picture, we appeal to the idea that an agent's action can only be captured by a propositionally specified reason if the agent knows the proposition in question.[30] If

---

[29] An anonymous referee voiced an objection to this idea, saying, "our whole penal system seems built on the falsity of the first claim—that we should punish only when it is, inter alia, for the reason the defendant did the crime. Our punishment regime, and nearly any one imaginable in a large society, is designed to punish at least some merely because they probably did it." One way to address this point (also noted by the referee) was that we could appeal to Moss' (forthcoming) idea that we can have probabilistic knowledge. This is one way to go and I should stress here that I'm trying to motivate a general approach, not some highly specific account. I did want the word 'matters' to operate as a kind of hedge above. One way that knowledge could matter is that we have a standard that says that we convict iff we know the defendant to be guilty. This would give use the desired result in Prisoners, but it would clash with the referee's observation about the role of probability in the criminal justice system. I think it's an interesting question whether the system that the referee defends is one that we should accept. (Part of this goes back to the earlier issue about objectivism about norms that I wanted to bracket.) We could say that knowledge matters in a different way—the difference in value for <Known Guilt, Punish> and <Unknown Guilt, Punish> tells us something about the values realized by outcomes—that they are sensitive to the presence or absence of knowledge. This evaluative claim doesn't tell us when it's right to convict. For that we need a norm that tells us how to respond to this value. We could try to vindicate the referee's observations by adopting a view on which these values and probabilities concerning what the agent knows will determine what the agent ought to do. In other words, we could make room for just conviction of defendants who are in fact innocent if we say that what we ought to do is maximize expected value as characterized by the gnostic. This would build on some ideas of Dutant (forthcoming) and Dutant and Fitelson (MS). I see a number of interesting avenues to pursue and I hope that readers don't assume that I'm assuming here that we need knowledge of the defendant's guilt to rightly punish. I assume that such knowledge is required for the ideal outcome to obtain.

[30] For defense, see Alvarez (2010), Hyman (2015), and Unger (1975).

we go this route, we can say that the reason that the law ought to assign different values to <Punish, Known Guilt> and <Punish, Unknown Guilt> is that it's only when the agent knows that the defendant did the deed that the agent's reason for imposing the harms could be the fact that the defendant did the deed. If the agent acts without this knowledge, the agent's reason could not be that the defendant did the deed. This, in turn, would be a misuse of the powers associated with the criminal process and should be regarded as a bad outcome in the eyes of the law.

On the second way of bringing knowledge into the picture, we don't need to appeal to the idea that it's only possible for the agent's reason to be that p if they know that p. Instead, we could appeal to the idea that the agent's reason couldn't be p if the agent knows they don't know whether p. If the agent knows or appears to know, say, that the defendant did some deed, their motive for acting could still be one that pertains to some offense that the law cares about. If, however, the agent neither knows nor appears to know that the defendant did some deed, the proper specification of the reasons for which they act wouldn't be that the defendant did some deed. As before, to impose the harms associated with punishment while conceding that the defendant might not have committed the offense in question would be a misuse of the powers associated with the criminal law. The law should see this as a bad outcome.

Because the law should care about the difference between (a) and (b), it ought to assign different values to imposing harms upon those we know (or seem to know) committed some offense and the imposition of harms upon those who we don't even seem to know committed that offense. The Gnostic Decision-Matrix reflects this. The Veritist Decision-Matrix does not. If you plug in bad values, it shouldn't be all that surprising that the norms from decision theory won't necessarily deliver good verdicts. If you plug in the right values, however, it won't be surprising that the verdicts will be less objectionable.

The takeaway from this is that we can see knowledge entering the picture by virtue of the connection between knowledge and motivating reasons. If readers prefer to do things with a rule and without decision tables, they can adopt a gnostic view on which the operative rule is Reasonable Conviction and then offer some gnostic account of what reasonable belief is.[31] If readers prefer to do things in terms of decision-theoretic norms and the promotion of value, it's clear that if the right decision matrix to use is anything like the Gnostic Decision-Matrix sketched here we don't maximize the relevant kind of expected value in Prisoners if we punish.

---

[31] Moss (forthcoming) notes that a problem with a knowledge account (or a justified belief account) of the standard of proof gives us only a plausible story about criminal law. It tells us nothing about civil cases that use a preponderance of evidence standard. We can generalize these proposals to the civil case in one of two ways. First, we might follow her lead in applying her notion of probabilistic knowledge (i.e., knowledge in which beliefs have probabilistic contents) to give us a knowledge-based account of the standard of proof in civil trials. Second, we might follow Blome-Tillman's (forthcoming) lead in arguing that we need knowledge in both cases but that the knowledge is easier to attain in civil trials because the stakes are lower. My aim here is to explain why the concept of knowledge should figure in the story and why the decision-theoretic argument we started with shouldn't convince us that high probability is sufficient for just conviction.

8 Is gnosticism necessary?

In part because I think we should consider Believe and Punish in tandem, I see Prisoners as another test case for our theories of epistemic justification and norms. The belief that the defendant was responsible for the assault rationalizes certain reactive attitudes. The belief, if justified, should justify these attitudes. If the thinker's grounds could not justify these reactive attitudes they could not justify the rationalizing belief. The evidence that we have in Prisoners doesn't justify the reactive attitudes. Statistical evidence justifies high confidence, but not outright belief. As Adler observed, "Mild resentment is never resentment caused by what one judges to be a serious offense directed toward oneself tempered by one's degree of uncertainty in that judgment" (2002: p. 217). Blame, like resentment, requires a commitment to truth that differs from mere high confidence. Thus, Prisoners shows us that views like the Lockean view that take the high probability of truth to be sufficient for justification or proper belief are inadequate. Such grounds justify high confidence without justifying the kind of untempered commitment to truth required by affective responses and reactive attitudes.

The Lockeans miss this because they see the grounds that justify high confidence as grounds that, *inter alia*, justify full belief. The same problem arises for some reliabilist views that, similarly, regard high probability of truth as the essential feature of justification.[32] If, as I've argued, Believe is false and this is a counterexample to the simple Lockean view of rational full belief, it is a counterexample to any reliabilist view that classifies lottery beliefs as justified. Most of us, I hope, would agree that the following thesis is false:

> Blame: It is appropriate to blame the defendant in Prisoners for assaulting the guard (and similar cases where the only evidence of guilt is statistical evidence).

When it's inappropriate to blame an agent, it might be because the agent cannot be held accountable or because no wrong was committed. If it's inappropriate to blame our defendant, it's not for these reasons. In our case, it is inappropriate to blame just because it is inappropriate to take it as settled that the defendant did the deed in question. In our case, we couldn't justify the complex attitude of taking the defendant to have done the deed and failing to blame or censure. This is just what the Lockean or the reliabilist is committed to if they accept Believe but reject Blame. On these views, the justified complex attitude would be expressed by saying, 'Look, you assaulted the guard but I cannot blame you for that'. There's no justification for this stance if you think the guard didn't deserve to be assaulted.

The Lockeans probably missed this because their focus has been on one kind of rationalizing relation, the relation between belief and behavior. They've neglected the rationalizing relation that holds between belief and emotion. An aversion to rain will

---

combine with full belief and result in a decision to take an umbrella. In the absence of a full belief, sufficiently high confidence will also do the trick. The same isn't true for our reactive attitudes and so couldn't be true for the full beliefs that are required for their rationalization.

In this final section, we consider alternatives to the gnostic view offered above to see if they can solve our puzzle. Our focus will be on Sosa's recent work and some work inspired by it. Much of what I say about knowledge and its value is consistent with Sosa's views, but there are some small differences that might matter.

In his discussions of judgment and belief, Sosa characterizes a kind of success and correctness in terms of truth or accuracy but it's clear that he doesn't think that success (so understood) represents the highest kind of epistemic good. He defends two theses in the course of defending his performance normativity framework:

> Success is better than failure.

> Success through competence is better than success by luck (2011: p. 63).

Sosa's proposals are about endeavors in general, not just epistemic endeavors that involve judgment and belief. When it comes to belief and judgment, success is understood as accurate belief and success through competence is apt belief. On Sosa's view, apt belief is a kind of knowledge, animal knowledge. When apt belief is aptly noted, a thinker has a further kind of reflective knowledge.

Thus, Sosa would agree with the veritists that true belief is better than false belief from the epistemic point of view and with the gnostics that knowledge is better than true but inapt belief.

Developed in one way, the view would differ from both the simple veritist view and the gnostic view in that it would incorporate two theses:

> Accuracy is Good-Making: the accuracy of a true belief is, from the epistemic point of view, a good-making feature.

> Aptness is Better-Making: the features of a belief that make it apt make it better, from the epistemic point of view, than accurate but inapt belief.

Some sophisticated veritists accept this much and accept a further claim about the value of accurate belief:

> Accuracy is Better than Nothing: the accuracy of a true belief is, from the epistemic point of view, a good-making feature that makes it better to hold an accurate belief about p than no belief at all.

According to Accuracy is Better than Nothing, the accurate belief is overall good, albeit a good that is inferior to apt belief. I think that some sophisticated veritists who accept the two central theses of the performance normative framework are sympathetic to Accuracy is Better than Nothing. I'll have more to say about these theses in a moment.

In explaining why we should reject the simple Veritist Decision-Matrix and Belief-Matrix, I didn't frame things in terms of the comparative value of success and success through competence. As it happens, I am quite sympathetic to the idea that knowledge is an accurate belief where the accuracy manifests the thinker's abilities, but the story

that I'd offer to explain Aptness is Better-Making differs from Sosa's in a few respects. First, my proposal doesn't appeal to the idea that belief or judgment involves any sort of endeavor or performance on the thinker's part. Thus, the success of the story doesn't turn on whether we should think of knowledge (or apt belief) as an achievement.[33] It also isn't part of my story that success should be understood in terms of accuracy. While I don't think that any inaccurate belief can do what beliefs are supposed to, I also don't think that every accurate belief does what beliefs are supposed to. On my account, the value of a belief turns on whether it can do what beliefs are supposed to do. The point, purpose, or function of belief is to provide us with reasons that consist of facts, putting them into our possession so that they can be our reasons for thinking things, feeling things, and doing things. A belief can do this iff it constitutes knowledge or iff it is apt. Thus, on my gnostic view, Aptness is Better-Making is true because Aptness is Good-Making is true and every inapt belief is taken to be bad from the epistemic point of view.

This proposal about the relationship between belief, knowledge, and our potential motivating reasons played an essential role in my gnostic solution to the puzzle. My proposal was that the law should recognize the superior value of knowledge and prefer the Gnostic Decision-Matrix to the Veritist Decision-Matrix on the grounds that the law should care about whether a jury's reason for convicting could be the fact that the defendant was involved. If not, the conviction should be seen as bad. If so, the conviction could be seen as good. The value of the conviction in the law's eyes depends upon what the jury could have known because this value turned on what the jury's reason for convicting could have been. Sosa takes a dim view of this approach to the value of knowledge:

> …one's action falls short if it is based on ostensible reasons that one does not know to be true. This is not because of the fact that a proposition can constitute your reason for X'ing only if it is something you know to be true. This is, I believe, at most a fairly superficial fact of English. But rather, there is a deeper, closely related truth here, which can be put in terms of one's rationale, of one's ostensible reasons, or of propositions adduced as reasons, or of stative reasons: i.e., beliefs on which one bases some further belief, or some choice or decision. The normative truth of interest is that if one acts based on a basis reason (or rationale), and if this reason is not something one knows to be true, then one's action falls short (2011: p. 46).

To illustrate this idea, he remarks:

> When someone flips a switch as a means to turning on a light, for example, he has an ostensible reason on which (in a broad sense) he bases his action, namely that flipping the switch is a means to turning on the light. Now, any action taken as a means to a further objective will of course fall short if it does not bring about that further objective. Moreover, it will still fall short if the objective is attained by a certain kind of luck: i.e., in a way that does not manifest the agent's competence (2011: p. 46).

---

[33] For criticism of the idea that knowledge is an achievement, see Hacker (2013).

If I understand Sosa's suggestion, it seems to be that the real reason that Aptness is Better-Making is due to the fact that an apt belief is an instance of the general phenomenon of success manifesting excellence of ability and that this connection between knowledge and propositionally specified reasons is not the explanatorily relevant aspect.

If we play up the link between knowing p and manifesting competence and play down the link between knowing p and ensuring that p is among the reasons that could be an individual's reasons for ϕ-ing, I worry that we'll struggle to meet the Enoch, Fisher, and Spectre challenge. To meet that challenge, we have to show that the law should assign values to our decision-matrix to show that the law should think it's better not to punish in Prisoners. Suppose we think of success as consisting of things like accurately representing guilty defendants as guilty and correctly imposing harms upon the guilty and not the innocent and think of success that manifests competence as doing such things knowingly. About this proposal we might ask two questions. First, should the law be convinced by the performance normativity argument for Aptness is Better-Making? If the law were moved by these kinds of considerations, what view should it take concerning Accuracy is Better than Nothing?

With respect to the first issue, it's not clear that the law should be moved by the kinds of considerations that Sosa offers in support of Aptness is Better-Making. It looks like Sosa and I might agree on this point: when a thinker ϕs, her reason for ϕ-ing could not have been that p if she knew she didn't know p. On my account, if the relevant action is voting to convict, convicting, or punishing, such acts aren't permitted unless done in the awareness that the defendant had committed some wrong because such acts aren't permitted unless they express blame. This awareness requires knowledge. The cases that fit this description will coincide with the cases where success manifests competence, but should we think that this feature is doing the explanatory work? One reason to think that it doesn't is that it seems to draw our attention to the wrong kind of thing, whether the jurors brought about the correct outcome or whether they brought it about excellently. It seems that the law should care about things like whether the punishments serve their intended function (i.e., expressing blame and leading to outcomes whereby it is just the guilty that suffer punishment) and not on further questions about whether those responsible for their distribution put on excellent performances. If this is right, it isn't clear that the performance normativity framework gives us the resources we need to justify the adoption of the decision-matrices that assign greater value to <Punish, Known Guilt> than <Punish, Unknown Guilt>.

With respect to the second issue, we need to know what values Sosa or a sophisticated veritist would assign to <Believe, Unknown Guilt> and <Punish, Unknown Guilt>. A sophisticated veritist who sees accurate or true belief as a kind of success might also see punishing the guilty as kind of success, in which case they might adopt a view that incorporates Accuracy is Better than Nothing and Aptness is Better-Making and offer us a third set of decision and belief matrices:

	Known Guilt	Unknown Guilt	Innocent
Punish	+1	+1 > v > −10	−10
Don't Punish	0	0	0

	Known Guilt	Unknown Guilt	Innocent
Believe	+1	+1 > v > −10	−10
Don't Believe	0	0	0

Because the sophisticated veritist sees accurate belief as a state that is always better to have than to lack, the decision-theoretic argument that the Lockeans used to try to justify Believe is back on the table. Even when we know that we could not hope for success through competence (e.g., when we're dealing with propositions that we know that we couldn't know), we could also know that owing to the likelihood of success or accuracy, the expected epistemic value of Believe could be positive. If so, the expected epistemic value of Believe would exceed that of Don't Believe since that option never realizes any value at all.

Once the argument for Believe is back on the table, the case for Punish is back on the table. The sophisticated veritist might try to find some way to resist that argument, but remember that the proponents of the performance normativity framework often try to motivate their view by asking us to consider analogies between epistemic and practical performances. If someone thinks that these analogies are helpful in thinking about practical and epistemic normativity, they might be disposed to think that the kinds of argument that supported Believe would also support Punish.

One nice feature of Sosa's own view is that he doesn't accept the sophisticated veritist view just sketched. While he accepts Accuracy is Good-Making (and Aptness is Better-Making), he doesn't think that accurate belief is better than no belief at all. As he might put it, his view is about the value of performances. A false belief is a bad performance (in a way) but the absence of belief is not any sort of performance at all. Since he doesn't compare the value of these performances to non-performances (or beliefs to the absence of belief), he doesn't defend the view that <Believe, Unknown Guilt> is better than <Don't Believe, Unknown Guilt>. As such, the decision-theoretic reasoning that is supposed to support Believe wouldn't move him.

Because the sophisticated veritist's value theory implies that Believe maximizes expected epistemic value, the sophisticated veritist can only resist the argument for Believe by defending norms that tell us that we should sometimes refrain from believing or doing that which we know would maximize the relevant kind of expected value. Because Sosa's value theory has no such implication, it might seem that Sosa would have an easier time blocking the arguments for Believe and Punish.

While Sosa does reject the Lockean view that says that it's always rational to fully believe when it's rational to have some sufficiently high degree of confidence, there still might be aspects of his view that make it difficult for him to reject Believe and

Punish. Because Sosa's account of epistemic goodness is an account of the comparative value of performances, it differs from the sophisticated veritist view insofar as it doesn't suggest that accurate but inapt belief is better than the absence of belief. (It's this feature of his view that allows him to accept Accuracy is Good-Making without any commitment to Accuracy is Better than Nothing.) While there's no clear commitment on Sosa's part to the premises of the arguments that support Believe and Punish, it's also not clear that Sosa's view gives us the tools we'd need to reject these theses and show where the arguments offered in their support went wrong.

One key point of difference between Sosa's account and the sophisticated veritist account just sketched is about whether it would be regrettable for a thinker to fail to add an accurate or apt belief to her stock of beliefs. One key point of agreement between these accounts is that they seem to accept this kind of conditional: if the thinker is to form a belief about whether p (because she aims to settle the question whether p), it would be better for her to believe accurately than inaccurately and better still to believe aptly than inaptly. What would Sosa say about cases where the thinker endeavors to settle the question, sees that she could not aptly affirm that p but also sees that there is very little risk of error? Here it is helpful to think about Sosa's approach to suspension of judgment:

> Suspension of judgment is an intentional double- omission, whereby one omits affirmation, whether positive or negative. Inherent to rational suspension is the assessment that affirmation is then too risky, which implies that, whether positive or negative, it would not then be apt, or at least the absence of assessment that it would be apt (2015: p. 83).

In situations where suspension of judgment is not mandatory, either belief or denial would be permitted. This is because, as Sosa puts it, suspension, and denial as part of a threefold choice. In cases like Prisoners, it certainly *seems* that affirmation wouldn't be too risky, not if we think of the relevant kind of risk as risk of *error*. The risk of error in this case is lower than in some cases of apt belief, so it would seem that even if the judgment in the defendant's guilt isn't apt or doesn't constitute knowledge, suspension wouldn't be mandatory precisely because the risk of inaccuracy is so low. If so, it looks as if Sosa's view might be like the Lockean's view insofar as both views seem to imply that it is *permissible* to believe once the probability of accuracy is sufficiently high even if it differs from the Lockean view about whether such beliefs are *required*. If so, it looks as if Sosa's view might support Believe.

The problematic case, from my point of view, is the case in which a thinker can see that there's very little chance that her belief would be inaccurate even though it's clear that her belief would be inapt. Sosa might try to meet these challenges in a few ways. First, he might say that the permissibility of belief turns on the risk of *inaptness*, not just *inaccuracy*. I fear that this way of going makes the standards governing belief, punishment, and blame too demanding. If a thinker aptly believes the defendant is guilty, shouldn't this be sufficient for belief, blame, and punishment? A belief can be

apt even if, given the thinker's evidence, it's highly unlikely to be apt.[34] Thus, I don't think we'd want to say that a thinker should suspend whenever there is high risk of forming an inapt belief because such situations are sometimes situations in which a thinker could nevertheless form an apt belief.

In trying to undermine the arguments for Punish and Believe, I've assumed that cases relevantly similar to lottery cases are cases of inapt belief. All of the arguments offered here assume that a lottery belief is inapt even if accurate. Sosa has contested this. He thinks that we *can* know lottery propositions (2015: 120). If Sosa is right about this, this would undo all of my arguments in this paper.

One reason that I wanted to discuss this puzzle in a discussion of Sosa's epistemology is that the case provided us with an interesting new perspective on lottery propositions. While this isn't true of every reader, I expect that many readers agree that Punish is wrong. The crucial intuition is not universally shared, but it is widely shared. If I'm wrong and we can know lottery propositions, how can we account for the intuitions that Punish and Blame are false? On its face, it seems that *if* the jury knew that the defendant was guilty in some case (and the evidence that constituted the basis of this knowledge was admissible), it should also seem appropriate to punish. And yet, it seems inappropriate to punish.

Sosa could say, quite rightly, that there might be further reasons not to punish that could explain why we shouldn't punish even when we know that a defendant is guilty. There might be some interesting reason why the law should treat certain kinds of knowledge as pieces of intelligence that we shouldn't act on. For my own part, I don't know what this story would look like. I worry that even if such a story could be told, we'd need a further story to handle these issues in non-legal contexts. For my own part, it seems that just as it would be inappropriate to punish in Prisoners, it would *also* be inappropriate for people to take up certain reactive attitudes towards the defendant. Again, it seems strange to think that it would be inappropriate to take up reactive attitudes towards someone if you truly knew that they did something truly awful. If we work from the assumption that we don't know lottery propositions, it seems we have lots of the tools available for vindicating these intuitions. If, however, we work from the assumption that we know lottery propositions, it seems we wouldn't have any of these tools. Thus, I see the intuitions that generate our initial puzzle as intuitions that provide further support for the position that we don't know lottery propositions. Readers who think that we can know such propositions should try to generate their own solution to the puzzle.[35]

---

[34] See Williamson (2011) for a defense of improbable knowing.

References

Adler, J. (2002). *Belief's Own Ethics*. Cambridge: MIT University Press.

Ahlstrom-Vij, K., & Dunn, J. (2014). A defence of epistemic consequentialism. *The Philosophical Quarterly*, *64*, 541–51.

Alvarez, M. (2010). *Kinds of reasons*. Oxford: Oxford University Press.

Apt, B. (2016). Do we know how to punish? New criminal law review. *In International and Interdisciplinary Journal*, *19*, 437–472.

Berker, S. (2013). The rejection of epistemic consequentialism. *Philosophical Issues*, *23*, 363–387.

Bird, A. (2007). Justified judging. *Philosophy and Phenomenological Research*, *74*, 81–110.

Blome-Tillmann, M. MS. The knowledge norm of legal liability.

Blome-Tillmann, M. (2015). Sensitivity, causality and statistical evidence in courts of law. *Thought*, *4*, 102–112.

Buchak, L. (2013). Belief, credence, and norms. *Philosophical Studies*, *169*, 1–27.

Carr, J. (2015). Epistemic expansions. *Res Philosophica*, *92L*, 217–236.

Christensen, D. (2004). *Putting logic in its place*. Oxford: Oxford University Press.

Colyvan, M., Ferson, S., & Regan, H. (2001). Is it a crime to belong to a reference class? *The Journal of Political Philosophy*, *9*, 168–81.

Comesana, J. (2009). What lottery problem for reliabilism? *Pacific Philosophical Quarterly*, *90*, 1–20.

Dietz, C. (Forthcoming). Reasons and factive emotions. *Philosophical Studies*.

Dorst, K. (Forthcoming). Lockeans maximize expected accuracy. *Mind*.

Dutant, J. (Forthcoming). Knowledge-based decision theory and the new evil demon problem. In J. Dutant & F. Dorsch (Eds.), *The new evil demon problem: New essays on knowledge, rationality, and justification*. Oxford: Oxford University Press.

Dutant, J., & Fitelson, B. MS. Knowledge centered epistemic utility theory.

Easwaran, K. (2016). Dr. Truthlove or: How I learned to stop worrying and love Bayesian probabilities. *Nous*, *50*, 816–33.

Easwaran, K., & Fitelson, B. (2015). Accuracy, coherence, and evidence. In T. S. Gendler & J. Hawthorne (Eds.), *Oxford studies in epistemology* (Vol. 5, pp. 61–96). Oxford: Oxford University Press.

Enoch, D., Fisher, T., & Spectre, L. (2012). Statistical evidence, sensitivity and the legal value of knowledge. *Philosophy and Public Affairs*, *40*, 197–224.

Firth, R. (1981). Epistemic merit, intrinsic and instrumental. *Proceedings and Addresses of the American Philosophical Association*, *55*, 5–23.

Foley, R. (2009). Belief, degrees of belief, and the Lockean thesis. In F. Huber & C. Schmidt-Petri (Eds.), *Degrees of belief* (pp. 37–47). Berlin: Springer.

Gibbons, J. (2013). *The norm of belief*. Oxford: Oxford University Press.

Goldman, A. (1986). *Epistemology and cognition*. Cambridge: Harvard University Press.

Gordon, R. (1987). *The structure of emotion*. Cambridge: Cambridge University Press.

Hacker, P. (2013). *The intellectual powers*. Hoboken: Wiley.

Harman, G. (1968). Knowledge, inference, and explanation. *American Philosophical Quarterly*, *5*, 164–73.

Horowitz, S. (2014). Immoderately rational. *Philosophical Studies*, *167*, 41–56.

Hyman, J. (2015). *Action, knowledge, and will*. Oxford: Oxford University Press.

Jenkins, C. (2007). Entitlement and rationality. *Synthese*, *157*, 25–45.

Joyce, J. (1998). A nonpragmatic vindication of probabilism. *Philosophy of Science*, *65*, 575–603.

Kelp, C. (Forthcoming). Lotteries and justification. *Synthese*.

Lempert, R. (1977). Modeling relevance. *Michigan Law Review*, *75*, 1021–57.

Littlejohn, C. (2012). *Justification and the truth-connection*. Cambridge: Cambridge University Press.

Littlejohn, C. (2015). Who cares what you accurately believe? *Philosophical Perspectives*, *29*, 217–48.

Littlejohn, C. (Forthcoming). Stop making sense? On a puzzle about rational belief. *Philosophy and Phenomenological Research*.

Littlejohn, C. (Forthcoming b). The right in the good: A defense of teleological non-consequentialism in epistemology. In K. Ahlstrom-Vij & J. Dunn (Eds.), *Epistemic consequentialism*. Oxford: Oxford University Press.

Lo, H. L. (2008). *A philosophy of evidence law*. Oxford: Oxford University Press.

Lord, E. (Forthcoming). What you're rationally required to do and what you ought to do (are the same thing!) *Mind*.

Lynch, M. (2004). *True to life*. Cambridge: MIT University Press.

McGlynn, A. (2013). Believing things unknown. *Nous*, *47*, 385–407.

Moss, S. (Forthcoming). *Probabilistic knowledge*. Oxford: Oxford University Press.

Nelkin, D. (2000). The lottery paradox, knowledge and rationality. *Philosophical Review*, *109*, 373–409.

Owens, D. (2013). Value and epistemic normativity. *Teorema*, *33*, 35–58.

Pettigrew, R. (2016). *Accuracy and the laws of credence*. Oxford: Oxford University Press.

Pritchard, D. (2016). Veritism and epistemic value. In H. Kornblith & B. McLaughlin (Eds.), *Alvin Goldman and his critics*. Hoboken: Wiley.

Pritchard, D. (Forthcoming b). Legal risk, legal evidence, and the arithmetic of criminal justice. *Jurisprudence*.

Redmayne, M. (2008). Exploring the proof paradoxes. *Legal Theory*, *14*, 281–309.

Ryan, S. (1991). The preface paradox. *Philosophical Studies*, *64*, 293–307.

Ryan, S. (1996). The epistemic virtues of consistency. *Synthese*, *109*, 121–41.

Schauer, F. (2003). *Profiles, probabilities, and stereotypes*. Cambridge: Harvard University Press.

Simon, R. (1969). Judges' translations of burdens of proof into statements of probability. In J. Kennelly & J. Chapman (Eds.), *The trial Lawyer's guide* (pp. 103–14). Mundelein, IL: Callaghan.

Smith, M. (2016). *Between probability and certainty*. Oxford: Oxford University Press.

Smith, M. (Forthcoming). When does evidence suffice for conviction?

Sosa, E. (2011). *Knowing full well*. Princeton: Princeton University Press.

Sosa, E. (2015). *Judgment and agency*. Oxford: Oxford University Press.

Staffel, J. (Forthcoming). Beliefs, buses, and lotteries: Why rational belief can't be stably high credence. *Philosophical Studies*.

Steele, K. MS. Crime, punishment, and 'individualized' evidence.

Sturgeon, S. (2008). Reason and the grain of belief. *Nous*, *42*, 139–165.

Sylvan, K. (Forthcoming). Veritism unswamped. *Mind*.

Talbot, B. (2014). Truth-promoting non-evidential reasons for belief. *Philosophical Studies*, *168*, 599–618.

Thomson, J. J. (1986). Liability and individualized evidence. *Law and Contemporary Problems*, *49*, 199–219.

Titelbaum, M. (2015). Rationality's fixed point (or: in defense of right reason) (Vol. 5). *Oxford studies in epistemology* (pp. 253–94). Oxford: Oxford University Press.

Tribe, L. (1971). Trial by mathematics: Precision and ritual in the legal process. *Harvard Law Review*, *84*, 1329–1393.

Unger, P. (1975). *Ignorance*. Oxford: Oxford University Press.

Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.

Williamson, T. (2011). Improbable knowing. In T. Dougherty (Ed.), *Evidentialism and its discontents* (pp. 147–64). Oxford: Oxford University Press.

Zimmerman, M. (2008). *Living with uncertainty*. Cambridge: Cambridge University Press.