Against Legal Probabilism

Martin Smith

> To hold that any finite degree of probability should constitute proof adequate to the conviction of an offender would in reality be to assert that, out of some finite number of persons accused, an innocent man should be sacrificed for the sake of punishing the rest – a proposition which is inconsistent with the humane spirit of our law as it is with the spirit of reason and justice.
>
> - Thomas Starkie, *A Practical Treatise of the Law of Evidence*, 1824, p573

Is it right to convict a person of a crime on the basis of purely statistical evidence? Many who have considered this question agree that it is not, posing a direct challenge to *legal probabilism* – the claim that the criminal standard of proof should be understood in terms of a high probability threshold. Some defenders of legal probabilism have, however, held their ground: Schoeman (1987) argues that there are no clear epistemic or moral problems with convictions based on purely statistical evidence, and speculates that our aversion to such convictions may be nothing more than an irrational bias. More recently, Hedden and Colyvan (2019, section VI) describe our reluctance to convict on the basis of purely statistical evidence as an 'intuition', but suggest that there may be no 'in principle' problem with such convictions (see also Papineau, forthcoming, section 6). In this paper, I argue that there is, in some cases, an in principle problem with a conviction based upon statistical evidence alone – namely, it commits us to a precedent which, if consistently followed through, could lead to the deliberate conviction of an innocent person. I conclude with some reflections on the idea that the criminal justice system should strive to maximise the accuracy of its verdicts – and the related idea that we should each strive to maximise the accuracy of our beliefs.

Key Words: Probability, criminal standard of proof, statistical evidence, accuracy

I PURE STATISTICAL EVIDENCE AND THE CRIMINAL STANDARD OF PROOF

Some theorists have been attracted to the idea that the criminal standard of proof corresponds to a high probability threshold in the order of 90 or 95%. On this picture, a defendant in a criminal trial should be convicted just in case the probability that he is guilty, given the admitted evidence, exceeds the threshold (see, for instance, McBaine, 1944, Cullison, 1969, section IIIA, McCauliff, 1982, *Brown v Bowen* 847, F.2d 342 (7th cir. 1988) at 345, 346, Schauer and Zeckhauser, 1996, section III, Walen, 2015, Hedden and Colyvan, 2019). My concern in this paper is with a *normative* version of this claim; irrespective of whether judges and juries actually *do* interpret the criminal standard of proof probabilistically (we will soon see some reasons for suspecting that they don't), the claim is that the criminal standard of proof *ought* to be interpreted in this way – that criminal trials ought to operate with a probabilistic standard of proof. Call this *legal probabilism*[1].

---

[1] As far as I'm aware, this term was first used, in reference to this kind of view, by Haack (2014, chap. 3). Haack offers a number of criticisms of the view, which are not directly related to the criticism I pursue here – though

One motivation for legal probabilism comes from the plausible idea that the criminal justice system should strive for *accuracy* in its verdicts – it should aim to convict the guilty and to avoid convicting the innocent. In one sense, setting a criminal standard of proof is a matter of striking a balance between these two sides of the accuracy goal: Set the standard too high and we make it too difficult to convict the guilty. Set the standard too low and we make it too easy to convict the innocent. What we seek is a standard that will enable us to convict as many criminals as possible, while maintaining a favourable *ratio* of true convictions to false – and, so the thought goes, it is a probabilistic standard that will best deliver on this (see Cullison, 1969, pp566-567, McCauliff, 1982, section IV). Suppose the evidential probabilities of guilt are 'well-calibrated' across cases in the sense that, out of all the cases in which the evidential probability of guilt is n%, the defendant is guilty in precisely n% of these cases. If the criminal standard of proof corresponds to a 95% probability threshold, such that every person who is convicted is at least 95% likely to be guilty, this assumption guarantees that at least 95% of these people will be guilty and at most 5% will be innocent. More generally, given the assumption of well calibration, the accuracy goal will be best served by a probabilistic standard of proof, where the probability threshold matches the lowest acceptable proportion of true convictions (see Hedden and Colyvan, 2019, p459)[2]

One significant challenge for legal probabilism comes from cases involving *pure statistical evidence*. Suppose an electronics store is struck by looters during a riot. 100 people carry televisions from the store, while the transaction record at the cash register indicates that only one television was legitimately purchased, though no receipt was issued. Suppose Joe is stopped by the police while carrying a television from the store and is promptly arrested and charged with theft. When his case comes to trial, the prosecution presents the evidence that Joe is one of the 100 people who carried a television from the store during the riot and that 99 of those people are guilty of theft. Joe concedes these facts, but maintains his innocence.

The evidence presented in this case makes it 99% likely that Joe is guilty – higher than any threshold that has been associated with the criminal standard of proof. According to legal probabilism, Joe should be convicted – this should, indeed, be regarded as an 'open and shut' case. Most who have considered such cases, however, hold that it would be wrong to base a conviction on purely statistical evidence of this kind. Further, it is not just theorists who appear to hold this view; an unwillingness to base findings of guilt or liability on evidence that is purely statistical in nature is, to some extent, evident in actual legal cases, and in psychological studies involving mock juries[3].

---

see pp74-75 for some discussion of pure statistical evidence (and for a response to Haack's criticisms see Hedden and Colyvan, 2019, section 3).

[2] There is, of course, no particular reason to accept the assumption of well calibration and, without it, a probabilistic standard of proof provides no assurance as to the proportion of true convictions. My argument against legal probabilism will in fact depend upon a series of cases in which the guilt probabilities *are* guaranteed to be well calibrated and, as such, accuracy-based support for a probabilistic standard of proof might be expected to be at its strongest.

[3] Relevant cases include *Day v Boston & Maine Railroad* 96 Me 207, 52A, 771 (1902) *Virginia & S.W. Ry. Co. v Hawk*, 160 F 348 (1908), *Evans v Ely*, 13 F.2d 62 (1926), *Sargent v Massachusetts Accident Co.* 307, Mass 246, 29 N.E. 2d 825 (1940), *Smith v Rapid Transit Inc.*, 317 Mass. 469, 58 N.E 2d 754 (1945), *Commercial Standards Insurance Co. v Gordon Transports Inc.*, 154 F.2d 390 (1946), *People v Collins*, 438 P.2d 33 (1968), *Guenther v Armstrong Rubber Co.* 406 F.2d 1315 (1969), *State v Carlson*, 267 N.W. 2d 170 (1978), *Baker v Bridgestone/Firestone Co.* 966 F. Supp 874 (1996), *United States v Shonubi*, 103, F.3d 1085 (1997), *R v Watters*, All ER (D) 1469 (2000), *Spencer v Baxter Intern. Inc* 163 F. Supp 2d 74 (2001). Relevant studies are reported in Wells (1992), Wright et al. (1996) and Niedermeier et al. (1999). I don't wish to overstate the claim that people are unwilling to base affirmative legal verdicts on evidence that is purely statistical – while it's clear that there

As emphasised above, however, legal probabilism is intended as a normative thesis and, as such, is not refuted by the mere fact that people are *reluctant* to convict on the basis of purely statistical evidence. As some legal probabilists have pointed out, allowing convictions on the basis of statistical evidence could well be in the interests of accuracy (Hedden and Colyvan, 2019, pp461-463, see also Eggleston, 1980, Brook, 1986, section IVC, Koehler and Shaviro, 1990). By convicting Joe, we are overwhelmingly likely to have added a true conviction to the tally, and by acquitting him we are overwhelmingly likely to have set a criminal free. Furthermore, many of the convictions that we *do* accept – such as convictions based upon eyewitness identification – may run a higher risk of error and serve the accuracy goal less well. Schoeman (1987) suggests that our reluctance to convict on the basis of purely statistical evidence may be nothing more than an irrational bias, while Hedden and Colyvan (2019, pp460, 464) describe it as an 'intuition', which may lack any principled motivation (see also Papineau, forthcoming, section 6). Even some of those who endorse the idea that affirmative legal verdicts should not be based upon purely statistical evidence – such as Williams (1980, p106) – describe the attitude as 'illogical'.

Theorists have highlighted several epistemic shortcomings of purely statistical evidence – it is insufficient for *knowledge* of a defendant's guilt (Littlejohn, forthcominga, forthcomingb) and it is incapable of making a guilty verdict *sensitive* (Enoch, Fisher and Spectre, 2012), *safe* (Pritchard, 2018) or *normically supported* (Smith, 2018, Gardiner, 2018). But even these observations need not trouble the legal probabilist – not without an explanation as to why the law should *care* about knowledge, sensitivity, safety or normic support, and care about it enough to compromise, for its sake, on the goal of accuracy (Enoch, Fisher, Spectre, 2012, section VA, Papineau, forthcoming, sections 8, 9). To adapt a thought experiment devised by Enoch, Fisher and Spectre, suppose we are presented with two criminal justice systems A and B and asked to choose a system under which our children will live. System A is more 'epistemically respectable' in that it places knowledge, sensitivity, safety or normic support requirements upon conviction and is unwilling to convict on the basis of purely statistical evidence. System B, however, is more accurate – perhaps it has the same ratio of true convictions to false, but more crimes end up leading to conviction, or perhaps the same proportion of crimes lead to conviction, but it has a more favourable ratio of true convictions to false, or perhaps it outperforms system A in both of these respects. Surely none of us would hesitate in choosing system B in this case – to choose system A would seem like a kind of 'epistemic fetishism' (Enoch, Fisher and Spectre, 2012, p213).

Other theorists have argued that there may be practical or policy reasons to block convictions based on purely statistical evidence – perhaps allowing such convictions could undermine public confidence in the criminal justice system (Nesson, 1979) or dampen the capacity of the system to deter crime (Enoch Fisher and Spectre, 2012, section VC). But these arguments rest upon rather speculative assumptions and, even if they are accepted, they pose no direct threat to legal probabilism unless we accept that avoiding these outcomes is more important than pursuing the accuracy goal. In the next section I will outline a somewhat different argument to the effect that Joe should not be convicted, closely related to an argument offered by Nunn (2015)[4]. Rather than pitting some other

---

is some tendency to this effect, the full range of relevant legal decisions and empirical findings presents a more complex and conflicted picture.

[4] In putting forward this argument I make no claim to novelty. In addition to Nunn, the argument is strongly suggested by Cohen (1977, §24) and Kaye (1979) who were amongst the first to discuss hypothetical cases with the structure of the riot example (see also Nesson, 1979). The argument may also be implicit in much earlier sources too, such as Starkie (1824, pp572-574).

consideration against the goal of accuracy, this argument invites closer scrutiny of the accuracy goal *itself*, and the ways in which it can be legitimately pursued.

II A SURFEIT OF GUILT

Return to the riot scenario. Suppose Joe is convicted of theft, on the grounds that he is one of the 100 people who carried a television from store and that 99 of these people committed theft. Now suppose that the owners of the electronics store produce CCTV footage that allows the police to reliably identify the other 99 people who carried televisions from the store during the riot. The evidence against each of these individuals is exactly the same as the evidence against Joe. Having convicted Joe on the strength of this evidence, we are effectively committed to convicting each of these individuals as well. To convict some, and not others, would be blatantly unjust. By convicting all 100 individuals, however, we are now guaranteed to have convicted an innocent – we have more convictions than we do crimes. To convict all 100 individuals is tantamount to deliberately convicting an innocent person.

Consider the following thought experiment – a familiar 'test case' for ethical theories: Suppose a community has been terrorised by a spate of crimes, and the resulting panic and anger has sparked violent protests, vigilantism and general civil disorder. Although we are unable to apprehend the true perpetrator we know that, by fabricating evidence and convicting an innocent person, we would be able to restore calm, order and security, and generally enhance the wellbeing of the community etc. The riot example presents a similar bargain – only it is the convicting of 99 guilty people that is on offer if we are willing to convict an innocent. Anyone who objects to the conviction of an innocent in the ethics test case would be very hard pressed to uphold the 100 convictions in the riot case. In both cases an innocent is sacrificed and, if anything, the consequent benefits are more modest in the latter than the former.

There are, of course, some notable differences between the two cases. In the ethics test case, as usually described, there is a particular, known individual we can choose to convict. In the present example, we have the choice of convicting some innocent individual or other, without knowing who the individual is. The former, we may say, is a matter of deliberately convicting an innocent *de re* while the latter is a matter of deliberately convicting an innocent *de dicto*. While this is worth pointing out, it has no obvious relevance for the permissibility of these actions. The intention element of criminal offences can be satisfied by intentions that are either de re or de dicto in the sense defined here. To satisfy the intention element of, say, murder or theft, it is sufficient that one intend to kill/take the property of *some person or other* (see for instance Duff, 2005, pp942-943)[5].

---

[5] Furthermore, we could easily adjust the ethics test case in such a way that it too involves only the de dicto deliberate conviction of an innocent. Suppose a leader orders that a randomly selected innocent person be made into a scapegoat for the crime spree (and convicted for the crimes) and decrees that he is never to be informed of this person's identity. While this may perhaps make it easier for the leader to live with his decision, it doesn't make it any more morally acceptable.

Another notable difference between the ethics test case and the riot case is that, in the former but not the latter, convicting the innocent involves deception, faking evidence etc. While such actions are obviously objectionable in their own right, it is clear that they are not the primary focus of the revulsion we feel towards this option. Thinking of the ethics test case, Rawls (1955, n8) remarks that the deliberate punishment of the innocent necessarily requires fraud. This may be true for the deliberate punishment of the innocent de re but, as the riot case shows, no fraud need be involved in deliberately punishing the innocent de dicto.

Consider the following argument:

(i)   It is permissible to convict Joe. (Assumption)
(ii) If it is permissible to convict Joe, then it would be permissible to convict each of the 100 individuals.
(iii) It is not permissible to convict each of the 100 individuals.
(iv) Therefore, it is not permissible to convict Joe. (Reductio)

Admittedly, this argument doesn't show that it is *never* permissible to base a conviction on purely statistical evidence. Not all convictions based upon statistical evidence will fit this pattern – that is, not all kinds of incriminating statistical evidence are such that they could simultaneously apply to each of a set of individuals, one or more of whom is guaranteed to be innocent. In any event, to refute legal probabilism, it is enough to describe a single case in which it is impermissible to convict a person on the basis of statistical evidence alone, even though it proves his guilt beyond the proposed probability threshold[6].

Here is another way of thinking about the present argument: When confronted with the 100 individuals who carried televisions from the store, we have but three options: (1) Acquit them all. (2) Convict them all. (3) Convict some and acquit others. (3) is ruled out on the grounds that any decision as to who to convict would be completely arbitrary – by stipulation the evidence against each individual is exactly the same. (2) is ruled out on the grounds that it involves deliberately convicting an innocent. This leaves (1) as the only alternative – every individual should be acquitted, including Joe, in spite of the fact that his guilt is highly probable, given the evidence[7].

A defender of legal probabilism could argue that, since any practicable legal system would be expected to mistakenly convict innocent people on occasion, we have no solid grounds for objecting to the conviction of all 100 defendants in this case. There is, of course, some risk that any given

---

[6] One possible way to insulate legal probabilism, in the face of such examples, is to argue that the purely statistical evidence in question should be deemed *inadmissible*, in which case the probability of guilt, conditional upon the admitted evidence, will remain low. I am inclined to think that this is not a promising strategy for a legal probabilist to pursue – as well as being rather implausible in its own right, such a standard of admissibility would be very difficult to square with the primary, accuracy-promoting rationale for adopting legal probabilism in the first place. I won't explore this any further here however. Thanks to Federico Picinali for discussion of this point.

[7] One might wonder whether the problem that I've levelled at the probabilistic interpretation of the criminal standard of proof is more general – applying perhaps to any interpretation on which the standard can be met by propositions that are less than certain. In reply, it's worth pointing out that the problem does not arise, at least in its present form, for any of the alternative interpretations alluded to at the end of section I – interpretations on which the criminal standard of proof requires knowledge, sensitivity, safety or normic support. These interpretations predict that all 100 defendants in the riot example should be acquitted – none has been proved guilty to the criminal standard. In fact, these interpretations predict, more generally, that the propositions proved to the criminal standard by a given body of evidence must form a consistent set. I won't discuss this further here.

As well as setting out a view on which the criminal standard of proof requires knowledge, Littlejohn (forthcomingb) also describes a view on which the criminal standard of proof is interpreted in terms of *probable knowledge* – a view on which a defendant in a criminal trial should be convicted just in case the probability that the fact finder *knows* him to be guilty, given the admitted evidence, exceeds a high threshold. While this interpretation delivers the desired result in the riot example, there may be other cases in which it would oblige the conviction of each of a set of defendants, one or more of whom is guaranteed to be innocent (depending upon how we understand the requirements for knowledge). If so, this view would also be subject to the problem I have outlined. I won't discuss this further here – for some discussion of this issue see Littlejohn (forthcomingb, section 3).

criminal conviction is mistaken and, as more and more convictions are handed down, the probability of mistaken convictions approaches certainty. And yet, reflecting upon these facts should hardly make us more willing to deliberately sacrifice an innocent in the riot scenario. There is a clear moral difference between tolerating a risk of mistaken convictions, as an inevitable consequence of a functioning criminal justice system, and deliberately convicting innocent people – obviously *not* an inevitable consequence of a functioning criminal justice system[8]. Further, if the existence of mistaken convictions makes it permissible, in the riot case, to deliberately convict an innocent, then it should have the same effect in the ethics test case – where it patently does not.

Drawing upon case law, Nunn (2015) has argued that convicting each of a set of defendants, when one or more is guaranteed to be innocent, would, in the United States, amount to an infringement of the defendants' due process rights. In *Smith v Groose* 205 F.3d 1045 (8th Cir. 2000) a murder conviction was quashed on the grounds that the state had used inconsistent theories in order to convict multiple defendants of the same crime, with a court finding that this constituted a violation of due process[9]. Significantly, the court appeared to endorse a more general principle to the effect that prosecutors cannot legitimately use 'inherently factually contradictory theories' in separate trials (at 1052, see also *Thompson v Calderon* 120 F.3d 1045 (9th Cir, 1997), section III). While this case is very different to the riot example, the general principle would appear to apply (in a surprisingly direct way) to this hypothetical scenario and to straightforwardly prohibit the conviction of all 100 defendants. In any case, I won't explore this further here.

In a way, the argument I have offered exposes the limits of the accuracy goal or, better, the limits on how the accuracy goal can be legitimately pursued. It is clear that convicting all 100 defendants will serve the goal of accuracy, as we have defined it. By doing this, we add 99 correct convictions to the tally and only one error – surely a very favourable ratio in anyone's estimation. By acquitting the 100 defendants, on the other hand, we forego this opportunity and resign ourselves to 99 criminals remaining unpunished. A criminal justice system that opted for the former would, all else equal, be more accurate than a criminal justice system that opted for the latter. And yet, in spite of its greater accuracy, it far from clear that the former system is the one that we would wish for our children.

---

[8] Another way to frame this issue is in terms of the distinction between *attacks* and *endangerments* (Duff, 2005) – between actions which deliberately cause harm and actions which merely risk causing harm. Attacks and endangerments are generally thought to have a different moral structure (Duff, 2005, section II) – in particular, when assessing the permissibility of an endangerment, it is crucial to attend to the level of risk imposed, and to the costs of further precautions. Stretching this terminology a little, we might say that, by setting a non-maximal standard of proof for criminal trials, we do *endanger*, to some extent, any innocent individuals who stand trial – but to deliberately convict an innocent constitutes an *attack* upon that person. For further related discussion, see Walen (2015, section IIA).

[9] The case concerns the double murder of Pauline and Earl Chambers, which took place in November 1983 in Kansas City, Missouri. Two individuals – Jon Smith and Michael Cunningham – were convicted, in separate trials, in connection with this crime, with each conviction resting heavily upon the testimony of an eyewitness Anthony Lytle. And yet, Lytle provided police with two completely different versions of events. In particular, the testimony that was used in Cunningham's trial was inconsistent with that used in Smith's and, indeed, inconsistent with Smith having any involvement in the crime – a fact which formed the basis of the subsequent successful appeal of his conviction. It is important to note, however, that appeals courts have not been uniform in their treatment of prosecutorial inconsistency, and there are other cases in which multiple convictions have been allowed to stand, in spite of being based upon what appear to be mutually inconsistent prosecutorial narratives (see Poulin, 2001).

The goal of accuracy, as defined, involves convicting as many criminals as possible while maintaining a favourable ratio of true convictions to false. In a way it is obvious that such a goal could, in principle, be furthered by deliberately convicting an innocent person – provided this served as a means to convicting a sufficiently large number of guilty people. As such, a criminal justice system that was willing to countenance such trade-offs may well be able to pursue accuracy more efficiently than one which was not. It is also obvious that the accuracy goal is intended, at least in part, to protect the innocent, and to deliberately harm innocents in its name seems as clear an example as we could imagine of the corruption of a worthy goal.

I have argued that our reluctance to convict Joe is no mere intuition – there is a principled argument to back it up. Hedden and Colyvan could point out that I am still appealing to intuitions however – the 'intuition' that it would be impermissible to convict all 100 defendants, or perhaps the 'intuition' that it would be impermissible to convict some of the individuals and acquit others. It is true, of course, that any argument must rest upon premises that are not themselves derived from anything else – and the argument will be unpersuasive for anyone who rejects them. I will end this section by suggesting a slightly different way of viewing things.

On one level, a criminal justice system is nothing more than a human invention, designed to serve particular ends. When we judge that it is impermissible to deliberately convict an innocent person in order to convict a number of guilty people, we are not attempting to 'intuit' the properties of some ideal system. Rather, it is better to see this as expressing a wish about *our own* system – the criminal justice system under which we live. What the argument establishes is something conditional: If we desire a system that would never condone the deliberate conviction of the innocent as a means to convicting the guilty, then we desire a system that does not operate with a purely probabilistic standard of proof. For many people, their faith in the criminal justice system is partly rooted in the belief that, so long as they abide by the law, the system will ensure, as far as is practical, that they are never convicted and subjected to punishment. It is vague, of course, exactly what this commitment amounts to – but no system that was willing to sacrifice the innocent in order to convict the guilty could claim to honour it.

III AGAINST EPISTEMIC PROBABILISM

I have argued that, in the riot case, it would be wrong to convict all 100 defendants, even though this would serve the aim of accuracy, as defined. While this is the primary conclusion of the paper, in this final section I will briefly pursue one line of thought that this conclusion naturally prompts. As many have observed, a criminal trial cannot be considered a 'pure' inquiry into the truth – rather, it serves several different aims and functions which have to be balanced against the goal of accuracy (for discussion see for instance Koehler and Shaviro, 1990, pp247-252, Ho, 2008, partic. chap. 2). It may be tempting to think about my conclusion in terms of this familiar narrative – as illustrating another way in which legal fact-finding is 'impure' and subject to demands which can impinge upon the accuracy goal. In spite of its appeal, I will argue that this thought is mistaken, and that the unfettered pursuit of accuracy is illegitimate even in a 'pure inquiry'.

As some will have doubtless observed, the dilemma that the riot case presents for a criminal justice system is reminiscent of that faced by the protagonist in the *lottery paradox*. Suppose I hold a single ticket – ticket #72 say – in a fair lottery with 99 losing tickets and a single winner. Suppose the lottery has been drawn but I've not yet heard the result. Surely, so the reasoning goes, I would be justified in believing that ticket #72 is not the winner – after all, given my evidence, this is 99% likely

to be true. But if I have justification for believing that ticket #72 has lost, then I must have justification for believing the same thing about each of the other tickets – as far as my evidence is concerned, all the tickets are on a par. According to the *principle of closure*, if I have justification for believing each of a series of propositions, I must also have justification for believing their conjunction. If we accept this principle, then I must have justification for believing something I know to be false – namely, that every ticket in the lottery has lost.

In the riot scenario, Joe is a member of a set of individuals, 99 of whom are guilty and one of whom is innocent. Surely, one might think, it would be permissible for a court to convict Joe – after all, given this evidence, it is 99% likely that he is guilty. But if it is permissible to convict Joe, it must be permissible to convict each of the 100 individuals – as far as the evidence is concerned, these individuals are on a par[10]. In a way, the 'legal' version of the paradox stops here. If I believe that ticket #1 has lost and that ticket #2 has lost and so on up to ticket #100 my attitude to the lottery outcome is, supposedly, not yet settled. According to the conventional view, there is still one further decision that I need to make – namely, whether to conjoin these beliefs and conclude that *every* ticket in the lottery has lost. Even if each of the 100 individual beliefs was justified, one could reject the principle of closure and deny that one would be justified in taking this extra step. Many have, in fact, taken the principle of closure to be pivotal to the lottery paradox, and have attempted to solve it in precisely this way (Kyburg, 1961, Schick, 1966, Derksen, 1978, Foley, 1993, chap. 4, 2009).

When it comes to the legal paradox, however, no such solution is in prospect. If the first individual has been convicted, and the second individual has been convicted right up to the hundredth individual, there is no *further* decision for the criminal justice system to make – no question as to whether we should now convict *every* individual. This has already been done – the acts, in a sense, 'conjoin' themselves. According to the closure-deniers, what we should do in the lottery scenario is believe, in the case of each ticket, that it has lost, while holding back from believing that all the tickets have lost. No corresponding advice is possible in the legal paradox – to convict in the case of each individual *just is* to convict all the individuals. If we judge that this is impermissible, and we hold that legal decisions cannot be made on an arbitrary basis, then only one solution to the 'paradox' is available – it is impermissible to convict any one of these individuals, including Joe. The equivalent solution to the lottery paradox is, of course, to insist that I lack justification for believing, of ticket #72 or any of the other tickets, that it has lost (Ryan, 1996, Nelkin, 2000, Smithies, 2012, Smith, 2010, 2016, chap. 3).

This strategy for resolving the lottery paradox puts us at odds with what we might call *epistemic probabilism* – the idea that one has justification for believing a proposition just in case the probability that it's true, given one's evidence, exceeds a high threshold. My evidence makes it 99% likely that ticket #72 has lost – higher than any threshold that has been associated with epistemic justification. One reason that epistemologists have been loath to adopt this solution to the lottery paradox is that it appears to be in tension with the idea that we should strive for *accuracy* in our beliefs. The claim that the ultimate aim of inquiry is to believe truths and avoid believing falsehoods is one that is very familiar. Almost as familiar is the following way of spelling it out: The ultimate aim of inquiry is to believe as many truths as possible while maintaining a favourable ratio of true beliefs

---

[10] The analogy with the lottery paradox is closest perhaps if we have a single judge deciding the 100 cases (which we could of course imagine). But it matters little if the cases are decided by different fact finders, provided they are each seen as implementing the policies of a single criminal justice system.

to false[11]. Many have held that securing epistemic justification is a matter of believing in a way that furthers this aim (see, for instance, Alston, 1985, Douven, 2008, Goldberg, 2018, 1.3, 1.4).

How then could I lack justification for believing, of any ticket, that it has lost, when forming this belief would clearly serve the interests of accuracy, so defined? By believing, of each ticket, that it has lost, I would add 99 true beliefs and only one false belief to the tally – surely a very favourable ratio in anyone's estimation. By refusing to form these beliefs, on the other hand, I would forego the opportunity to believe 99 truths. To insist that I lack justification for these beliefs, one might think, is tantamount to severing the connection between justification and accuracy. The considerations of the previous section, however, suggest that such a thought is too quick.

Suppose an omniscient being known as the 'truth fairy' offers to grant me a wealth of true, knowledgeable beliefs, and to correct a number of mistaken and misguided beliefs that I currently hold. Suppose that all she asks in return is that I bring myself to believe a proposition which I currently know to be false – say, that 1997 was a leap year. This is a variant on an example originally described by Jenkins (2008, see also Jenkins and Elstein, forthcoming) and used to put pressure on *epistemic consequentialism* – the view that a belief is justified whenever it serves to maximise epistemic value (for relevant discussion see Berker, 2013). Whatever we think of the truth fairy's offer, it doesn't make it the case that I suddenly have *epistemic justification* for believing that 1997 was a leap year. In spite of the epistemic benefits that such a belief would bring, I have no more justification for believing this proposition than I did before[12]. Here is an example of a belief which would clearly further the accuracy goal and which lacks justification nevertheless. This doesn't mean that there is anything suspect about the accuracy goal per se – even a legitimate goal can be pursued in illegitimate ways. Whichever means of pursuing the accuracy goal are conducive to justification, deliberately believing a falsehood is not amongst them.

In the lottery example, I face a choice not unlike that offered by the truth fairy – at the cost of deliberately adding a falsehood to my stock of beliefs, I am able to add 99 truths. Anyone who denies that I would be justified, in the truth fairy case, in believing that 1997 was a leap year would be hard pressed to maintain, on accuracy-seeking grounds, that I would be justified in forming these 100 beliefs. In both cases, I would come to deliberately believe something false and, if anything, the former would further the accuracy goal more effectively than the latter[13].

---

[11] This way of stating the accuracy goal for beliefs assumes that there is some principled way to 'count up' the number of truths and falsehoods that a subject believes or, at least, to meaningfully compare the sizes of these two sets. This assumption is undoubtedly problematic – see, for instance, Treanor, 2019, section 3 – but I put qualms about it to one side here. It is interesting to observe that this problem does not arise (in anything like as pressing a form) when stating the accuracy goal for criminal verdicts.

[12] In the original example, instead of asking that I come to believe a proposition which I currently know to be false, the truth fairy asks that I come to believe a proposition for which I have no relevant evidence one way or the other.

[13] In the truth fairy example, there is a particular false proposition which we are asked to believe – to do so would be to deliberately believe a falsehood *de re*. In the lottery example, we have the choice of believing some false proposition or other, without knowing which proposition it is – to take this choice would be to deliberately believe a falsehood *de dicto* (for related discussion see Smith, 2014). Once again, it is difficult to see why this difference should be of any relevance when it comes to the justificatory status of these respective choices. The onus, at any rate, would be on one who wished to suggest otherwise.

In this paper I have elaborated upon one standard objection to legal probabilism – the objection from pure statistical evidence. Rather than simply taking it for granted that one should not, in a case like the riot scenario, base a conviction on evidence that is purely statistical, this claim can be derived from the principle that one should never deliberately convict the innocent as a means to convicting the guilty. While there is a link between a probabilistic standard of proof and the goal of accuracy, I have argued that an increase in accuracy, if it comes via a willingness to deliberately convict innocent people, is of no credit to a criminal justice system. Relatedly, I have argued that an increase in accuracy, if it comes via a willingness to deliberately believe falsehoods, is of no credit to a believer.

References

Alston, W. (1985) 'Concepts of epistemic justification' *Monist* v68(1), pp57-89

Berker, S. (2013) 'The rejection of epistemic consequentialism' *Philosophical Issues*, v23, pp363-387

Brook, J. (1985) 'The use of statistical evidence of identification in civil litigation: Well-worn hypotheticals, real cases and controversy' *St Louis University Law Journal* v29, pp293-352

Cohen, L. (1977) *The Probable and the Provable* (Aldershot: Gregg Revivals)

Cullison, A. (1969) 'Probability analysis of judicial fact-finding' *University of Toledo Law Review* v1, pp538-598

Derksen, A. (1978) 'The alleged lottery paradox resolved' *American Philosophical Quarterly* v15(1), pp67-74

Douven, I. (2008) 'The lottery paradox and our epistemic goal' *Pacific Philosophical Quarterly* v89(2), pp204-225

Duff, A. (2005) 'Criminalizing endangerment' *Louisiana Law Review* v65(3), pp941-965

Eggleston, R. (1980) 'The probability debate' *Criminal Law Review* pp678-688

Elstein, D. and Jenkins, C. (forthcoming) 'The truth fairy and indirect epistemic consequentialism' in Petersen, N. and Graham, P. *Epistemic Entitlement* (Oxford: Oxford University Press)

Enoch, D., Fisher, T. and Spectre, L. (2012) 'Statistical evidence, sensitivity and the legal value of knowledge' *Philosophy and Public Affairs* v40(3), pp197-224

Foley, R. (1993) *Working Without a Net* (Oxford: Oxford University Press)

Foley, R. (2009) 'Beliefs, degrees of belief and the Lockean thesis' in Huber, F. & Christoph Schmidt-Petri, C. eds. *Degrees of Belief* (Dordrecht: Springer)

Gardiner, G. (2018) 'Legal burdens of proof and statistical evidence' in Coady, D. and Chase, J. eds. *Routledge Handbook of Applied Epistemology* (New York: Routledge)

Goldberg, S. (2018) *To the Best of Our Knowledge: Social Expectations and Epistemic Normativity* (Oxford: Oxford University Press)

Haack, S. (2014) *Evidence Matters: Science, Proof and Truth in the Law* (Cambridge: Cambridge University Press)

Hedden, B. and Colyvan, M. (2019) 'Legal probabilism: A qualified defence' *Journal of Political Philosophy* v27(4), pp448-468

Ho, H. (2008) *A Philosophy of Evidence Law: Justice in the Search for Truth* (Oxford: Oxford University Press)

Jenkins, C. (2007) 'Entitlement and rationality' *Synthese* v157(1), pp25-45

Kaplan, J. (1968) 'Decision theory and the factfinding process' *Stanford Law Review* v20(6), pp1065-1092

Koehler, J. and Shaviro, D. (1990) 'Veridical verdicts: Increasing verdict accuracy through the use of overtly probabilistic evidence' *Cornell Law Review* v75(2), pp247-279

Kyburg, H. (1961) *Probability and the Logic of Rational Belief* (Middleton: Wesleyan University Press)

Littlejohn, C. (forthcominga) 'Truth, knowledge and the standard of proof in criminal law' *Synthese*

Littlejohn, C. (forthcomingb) 'Justified belief and just conviction' in Robson, J. and Hoskins, Z. eds. *Truth and Trial* (Abingdon: Routledge)

McBaine, J. (1944) 'Burden of proof: Degrees of belief' *California Law Review* v32(3), pp242-268

McCauliff, C. (1982) 'Burdens of proof: Degrees of belief, quanta of evidence or constitutional guarantees?' *Vanderbilt Law Review* v35, pp1293-1335

Nelkin, D. (2000) 'The lottery paradox, knowledge and rationality' *Philosophical Review* v109(3), pp373-409

Nesson, C. (1979) 'Reasonable doubt and permissive inferences: The value of complexity' *Harvard Law Review* v92(6), pp1187-1225

Niedermeier, K., Kerr, N. and Messé, L. (1999) 'Jurors use of naked statistical evidence: Exploring bases and implications of the Wells effect' *Journal of Personality and Social Psychology* v76(4), pp533-542

Nunn, G. (2015) 'The incompatibility of due process and naked statistical evidence' *Vanderbilt Law Review* v68(5), pp1407-1433

Papineau, D. (forthcoming) 'The disvalue of knowledge' *Synthese*

Poulin, A. (2001) 'Prosecutorial inconsistency, estoppel and due process: Making the prosecution get its story straight' *California Law Review* v89(5), pp1423-1478

Pritchard, D. (2018) 'Legal risk, legal evidence and the arithmetic of criminal justice' *Jurisprudence* v9(1), pp108-119

Rawls, J. (1955) 'Two concepts of rules' *Philosophical Review* v64(1), pp3-32

Ryan, S. (1996) 'The epistemic virtues of consistency' *Synthese* v109(2), pp121-141

Schauer, F. and Zeckhauser, R. (1996) 'On the degree of confidence for adverse decisions' *Journal of Legal Studies* v25(1), pp27-52

Schick, F. (1966) 'Consistency' *Philosophical Review* v75(4), pp467-495

Schoeman, F. (1987) 'Statistical vs direct evidence' *Noûs* v21(2), pp179-198

Smith, M. (2010) 'What else justification could be' *Noûs* v44(1), pp10-31

Smith, M. (2014) 'Knowledge, justification and normative coincidence' *Philosophy and Phenomenological Research* v89(2), pp273-295

Smith, M. (2016) *Between Probability and Certainty: What Justifies Belief* (Oxford: Oxford University Press)

Smith, M. (2018) 'When does evidence suffice for conviction?' *Mind* v127(508), pp1193-1218

Smithies, D. (2012) 'The normative role of knowledge' *Noûs* v46(2), pp265-288

Starkie, T. (1824) *A Practical Treatise of the Law of Evidence* (London: J. and W.T. Clarke), page references are for the *Seventh American Edition* (Philadelphia: T and J.W. Johnson, 1842)

Treanor, N. (2019) 'The proper work of the intellect' *Journal of the American Philosophical Association* v5(1), pp22-40

Tribe, L. (1971) 'Trial by mathematics: Precision and ritual in the legal process' *Harvard Law Review* v84(6), pp1329-1393

Williams, G. (1980) 'A short rejoinder' *Criminal Law Review* v103

Walen, A. (2015) 'Proof beyond a reasonable doubt: A balanced retributive account' *Louisiana Law Review* v76(2), pp354-446

Wells, G. (1992) 'Naked statistical evidence of liability: Is subjective probability enough?' *Journal of Personality and Social Psychology* v62(5), pp739-752

Wright, E., Maceachern, L., Stoffer, E. and Macdonald, N. (1996) 'Factors affecting the use of naked statistical evidence of liability' *Journal of Social Psychology* v136(6), pp677-688