

# Rehabilitating Statistical Evidence

LEWIS ROSS 

*University of St. Andrews*

Recently, the practice of deciding legal cases on purely statistical evidence has been widely criticised. Many feel uncomfortable with finding someone guilty on the basis of bare probabilities, even though the chance of error might be stupendously small. This is an important issue: with the rise of DNA profiling, courts are increasingly faced with purely statistical evidence. A prominent line of argument—endorsed by Blome-Tillmann 2017; Smith 2018; and Littlejohn 2018 rejects the use of such evidence by appealing to epistemic norms that apply to *individual inquirers*. My aim in this paper is to rehabilitate purely statistical evidence by arguing that, given the broader aims of legal systems, there are scenarios in which relying on such evidence is appropriate. Along the way I explain why popular arguments appealing to individual epistemic norms to reject legal reliance on bare statistics are unconvincing, by showing that courts and individuals face different epistemic predicaments (in short, individuals can *hedge* when confronted with statistical evidence, while legal tribunals cannot). I also correct some misconceptions about legal practice that have found their way into the recent literature.

## 1. Overview

Recently, the practice of deciding legal cases on purely statistical evidence has been widely criticised.<sup>1</sup> Many feel uncomfortable with finding someone guilty on the basis of bare probabilities, even though the chance of error might be stupendously small. This is an important issue: with the rise of DNA profiling, courts are increasingly faced with purely statistical evidence. In a series of prominent papers, various philosophers appeal to epistemic norms governing individual inquirers to argue that legal verdicts should never be based on statistical evidence alone. My project is to show that not only have recent discussions mischaracterised how the law actually treats purely statistical evidence, but also argue that there are cases in which relying on such evidence is justifiable. In developing these claims, I demonstrate why we cannot solve jurisprudential questions by simply appealing to the norms found in theorising about individual epistemic agents. Rather, settling issues of legal theory requires us to look further afield, taking into the account the unique tasks and challenges that face legal systems.<sup>2</sup>

This argument will be developed over five sections. In §1, I present the intuitive case against purely statistical evidence, explain why it creates a jurisprudential puzzle, and

---

<sup>1</sup> For an overview see: Gardiner (2018). For specific arguments, Smith (2018) Littlejohn (2018) Blome-Tillman (2015; 2017) Di Bello (2019) Thomson (1986) Enoch et al (2012) provide representative examples.

<sup>2</sup> This paper focuses on the common law and will not discuss civil law traditions. However much of what I say is relevant to any legal system.

separate out different responses to this puzzle. In §2, I introduce key concepts of evidence law necessary to fully appreciate legal treatment of statistical evidence. In §3, I demonstrate why DNA profiling is properly characterised as statistical evidence before showing that—contrary to recent assertions—legal systems do not generally prohibit basing verdicts on bare statistics. In §4, I introduce and reject a range of arguments *against* the use of purely statistical evidence which appeal to epistemic norms that govern individual inquirers. These arguments overlook the fact that individuals, unlike courts, can *hedge* in response to statistics. Finally, in §5, I suggest that our qualms about purely statistical evidence stem from the fact that it makes the possibility of error salient. However, because managing the possibility of error in different contexts is a delicate yet necessary task that every fallible legal system must fulfil, I close by arguing that the best solution to the puzzle of statistical evidence takes each case on its individual merits, rather than entertaining any general prohibition against such evidence. One upshot is that we might reasonably convict in cases involving DNA evidence but not in some of the other cases that have particularly discomfited critics of statistical evidence.

## 2. The puzzle of purely statistical evidence

Below are two cases—taken from the criminal and civil domain respectively—standardly used to motivate the idea that there is something amiss about legal reliance on purely statistical evidence. (I'll sometimes call it 'PSE' for short).<sup>3</sup>

PRISONERS: 100 prisoners are exercising in the prison yard. 99 of them attack the guard, putting into action a plan that the 100<sup>th</sup> prisoner knew nothing about. The 100<sup>th</sup> 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.<sup>4</sup>

BLUE BUS: A bus causes injury to a pedestrian, but it is not known which company the bus belongs to. On the route where the accident occurred, the Blue Bus Company runs 75% of the buses. There is no further information that we can use to settle the question of which company the bus belongs to.

Many have the intuition that it would not be permissible to sanction in these cases. In the psychological literature, this phenomenon is called the 'Wells' effect' after an influential study of juror reluctance to assign liability on the basis of bare statistics.<sup>5</sup> These patterns of intuition give rise to the *puzzle* of purely statistical evidence because, given what we know about the fallibility of other types of evidence such as eye-witness testimony, statistical evidence will often be *more truth-conducive* than other sorts of evidence on which we routinely base legal verdicts. The issue, then, is the following: how concerned should we be about the intuitive reluctance many have about basing legal verdicts on PSE? Should we accept these intuitions and construct theories that vindicate them, or should we ignore them and focus on the fact that such evidence can be very reliable?

<sup>3</sup> Here, and throughout, evidence is used in the sense familiar to the legal system rather than as a philosophers' term of art. Thus, evidence refers to that which is adduced in the context of a trial to support or rebut some contention. It does not refer to any doxastic or epistemic state.

<sup>4</sup> Wording taken from Redmayne (2008).

<sup>5</sup> See Wells (1992).

Schematically, there are three different approaches to PSE. (In the following, what counts as ‘strong’ probabilistic support will vary depending on the standard of proof being used).

- Always:** It is always acceptable to base a legal decision on purely statistical evidence, so long as it provides strong probabilistic support to that verdict.
- Sometimes:** Sometimes it is acceptable to base a legal decision on purely statistical evidence when it provides strong probabilistic support to that verdict.
- Never:** It is always wrong to base a legal decision on purely statistical evidence even when it provides strong probabilistic support to that verdict.

We should hold at the outset that **Always** is incorrect. In some scenarios, there are specific moral reasons to avoid appealing to statistics in legal inquiry. One example might be the use of crime statistics about certain demographics. For instance, even if it were the case that 99.9% of all game console thefts were carried out by women, this shouldn’t be enough to convict some individual woman if, for example, a games console is stolen in a situation where only one man or one woman could have been involved. The repeated use of such statistics would, I suggest, amount to an *oppressive* pattern against certain demographic groups.<sup>6</sup> However, even though the use of statistics can constitute oppression in some cases, this worry does not generalise. For instance, in PRISONERS, a group of prisoners who happen to be exercising at the time of collective disorder does not constitute a stable demographic group. This suggests that there are different strategies for vindicating the intuitions we have about PSE cases.

One approach is to identify *specific features* of certain cases to explain why those particular types of case should not be settled by purely statistical evidence. For instance, one worry about PRISONERS might be that, if all 100 prisoners were captured, we would knowingly convict one of them wrongly. A rather different approach is to argue that *generic features* of all PSE cases mean that such evidence is not a fitting basis for a legal decision in general. The latter approach has been the most popular strategy, and has been used to motivate **Never**. This view will be my target—I do not think that the generic features shared in common by PSE cases suffices to justify a general prohibition against relying on bare statistics. In this sense, we should not draw an inference from our aversion to some cases involving PSE to every case involving PSE. Rather, in this paper, I will be advocating for **Sometimes** by arguing that different uses of PSE invoke different issues. On balance, sometimes we should allow that a legal decision can be based on purely statistical evidence. Furthermore, I will claim that these cases are relatively common in the law.

### 3. Evidence law: essential concepts

Before getting into the weeds, we must first briefly introduce three key concepts of evidence law: the Standard of Proof; the Evidential Burden; and Evidential Admissibility.

<sup>6</sup> See Mogensen (2019) on how racial profiling can constitute oppression. There is a growing literature on what we should think about such demographically targeted statistics. An important debate is whether forming beliefs on the basis of racial profiling is a *moral wrong* (see e.g. Basu 2019), an *epistemic wrong* (see e.g. Bolinger 2018) or something else. Here, I will be leaving this debate aside—I will be not be dealing with examples of statistical evidence that targets any particular demographic group.

Extant discussions have tended to be light on legal details, only focusing on the first of these. However, a cursory appreciation of other aspects of evidence law is essential to correctly understand the current place of statistical evidence in the law and ask normative questions about what its proper role should be.

### 3.1. *The Standard of Proof*

A central part of evidence law—familiar to most readers and well-trodden in recent philosophical literature—is the Standard of Proof ('SoP').

The SoP is the standard applied to determine whether a body of evidence provides strong enough support to render a specified legal (or quasilegal) judgement appropriate. There are various different SoPs which apply to different types of legal judgement. Everyone knows the 'beyond a reasonable doubt' ('BRD') standard that applies to verdicts in criminal cases. Most academics will also recognise the 'balance of probabilities' ('BoP') standard that applies to verdicts in civil cases (e.g. contractual disputes). BRD and BoP are not the only standards found in legal systems. Various other formulations are adopted by different (quasi)legal institutions for various purposes, imposing more and less exacting standards.<sup>7</sup> SoPs often apply outside the context of trial verdicts; for example, prosecutors are governed by a SoP when deciding whether to charge a suspect with a crime. The way in which different standards of proof have been formulated and interpreted has changed significantly over time.<sup>8</sup> For instance, the standard for criminal conviction was once glossed as 'moral certainty', a fallibilist contrast to the notion of 'mathematical' or 'metaphysical' certainty that is immune to sceptical doubt. The way in which current standards have been interpreted has also seen significant variation—for instance, attempts to clarify the BRD by quantification (e.g. a .95 credence) or by conflating it with the standards used in non-legal inquiry (e.g. the standards used to make an important decision in one's life) have waxed and waned in popularity over time.

To be explicit about how these standards work: the BRD standard, for example, dictates that 'a court should only issue a finding of criminal guilt iff the entirety of the evidence adduced supports the guilt of the accused beyond a reasonable doubt!' Judging whether or not a body of evidence meets the applicable SoP is left to the relevant *fact-finder*. In the context of a trial verdict, the fact-finder will be a judge or a jury. Deciding whether evidence meets the SoP is *not* a technical judgement in the sense of requiring the fact-finder to interpret the law. Rather, the role of the fact-finder is to use their everyday facility in discerning the plausible from the implausible to make a judgement about a body of evidence which they have had the opportunity to consider first-hand. Fact-finders, so long as they stay with bounds of reasonableness, have a broad discretion in judging whether or not the evidence meets the relevant SoP.<sup>9</sup>

<sup>7</sup> E.g. 'reasonable suspicion' (often governing police searches) and 'reasonable degree of likelihood' (used to assess UK asylum claims) are standards less demanding than the BoP. Other standards—e.g. 'clear and convincing evidence' (sometimes used in civil cases regarding children's welfare or psychiatric assessment) fall between the civil and criminal standards.

<sup>8</sup> See Laudan (2006) and Roth (2010) for discussion.

<sup>9</sup> This is why appellate courts are reluctant to overturn first-instance judgements about whether the evidence satisfies the SoP solely on the grounds that the appellate court disagrees with the original judgement.

### 3.2. Sufficient evidence: discharging the ‘evidential burden’

A widely overlooked yet crucial legal concept is that of leading *sufficient evidence* and discharging the evidential burden.

In contrast to the non-technical task of the fact-finder to decide whether a body of evidence meets whatever SoP is in play, the doctrine of sufficient evidence is a matter of *law* regulated by the judge. (Note: judges can be both legal arbiters and fact-finders—these are distinct roles).<sup>10</sup> The doctrine of sufficient evidence places a legal burden—often called the ‘*evidential burden*’—on the party aiming to establish some claim. This burden acts like a filter that weeds out unacceptably weak cases. Specifically, it means that the party aiming to establish some contention must adduce sufficient evidence in favour of that claim before it can legitimately be put to the fact-finder (who will then, taking into account *all* of the evidence, decide whether enough has been done to prove the contention on the relevant standard of proof).<sup>11</sup> For instance, a very minimal construal of sufficient evidence in the context of a criminal case is that *some* evidence must be adduced which: (i) suggests that a crime has been committed, and (ii) suggests that the accused is the perpetrator. If this burden is not discharged, then the court does not move into the deliberative stage of considering the evidence against the relevant SoP. For instance, to continue using criminal procedure as an example, a failure to discharge the evidential burden will force the judge to hold that the accused has ‘no case to answer’ and bring proceedings to a close. Some jurisdictions have more heavyweight conceptions of sufficient evidence than others. One rationale for stricter conceptions is to mitigate the possibility of miscarriages of justice by imposing more judicial oversight on cases before they are left to the jury.<sup>12</sup>

Crucially, sufficiency of evidence is not about the *quality* of evidence. The conditions under which evidence counts as sufficient is a technical matter of law for the judge to decide, whilst quality of evidence is a matter for the fact-finder. For example, it might happen that the prosecution in a criminal case adduces evidence that is rightly put before the jury, only for the jury to rationally find the evidence unpersuasive. This is entirely consistent; sufficiency of evidence in a criminal jury trial concerns whether enough evidence has been adduced to warrant putting the case to the jury in the first place, not whether enough has been done, *ultima facie*, to persuade the jury beyond a reasonable doubt.

### 3.3. Admissibility of evidence

A third concept of evidence law will be familiar to readers so we need only introduce it very briefly. This is the idea of *evidential admissibility*. Rules governing admissibility determine when certain types of otherwise relevant evidence must, as a technical matter of law, be excluded. There are a variety of different rationales behind different rules: for

<sup>10</sup> To illustrate: in a criminal *jury* trial the judge will be the legal arbiter but not the fact-finder. In a non-jury criminal case, the judge will be the legal arbiter throughout the trial *and then* play the role of the fact-finder in deciding whether or not the accused is guilty.

<sup>11</sup> There are differences in how the burden to prove facts are allocated in criminal and civil trials, but these nuances need not detain us for our purposes.

<sup>12</sup> An interesting example is found in the Scots’ law requirement of *corroboration*: in Scottish criminal trials, crucial facts must be corroborated by at least two independent sources of evidence to discharge the evidential burden. This further exemplifies the distinction between sufficiency of evidence and the SoP; e.g. one might suppose that one individual source of evidence (e.g. an eyewitness account) could conceivably establish guilt beyond a reasonable doubt, but due to lack of corroboration this—in a Scottish criminal trial—might fail to constitute sufficient evidence.

example, some rules are in place to discourage improprieties at other stages of the judicial process (for example, prohibiting evidence acquired using coercion during police interview) whilst other rules preclude evidence that is perceived as unreliable or irrelevant (for example, well-known rules about hearsay evidence).

When evidence should be excluded is a matter of law, determined by the judge. Evidence that is ruled inadmissible and excluded does not count towards discharging the evidential burden, and is not considered by the fact-finder in determining whether or not the applicable SoP has been met.

#### 4. Statistical evidence in the law: a closer look

Recent discussions create the impression that legal systems are generally adverse to purely statistical evidence. One influential paper asserts outright that such evidence is *inadmissible*.<sup>13</sup> Other philosophers do not locate the problem in admissibility, but nonetheless suggest that evidence law is unfavourable to bare statistics.<sup>14</sup> These are empirical assertions about the common law. Not only are they interesting in their own right, but, if they were true, they would lend credence to certain normative views. This is because it would suggest that **Never** is the orthodox position and imply that those hostile to PSE are simply seeking to vindicate rather than revise existing legal practice. (Indeed, Blome-Tillmann 2017 calls wholesale antipathy to PSE in the law *conservative* and opposing views *revisionist*).

Against this characterisation I show that close inspection reveals no general prohibition against PSE in the law. Rather, the rise of cases involving DNA profiling show that PSE is playing an ever more prominent role. Furthermore, brief reflection will show that the most promising way to reject PSE appeals to the idea of *sufficient* evidence rather than, as some have suggested, a rule of admissibility.

##### 4.1. DNA evidence is statistical evidence

In recent years the vast majority of reported legal cases involving PSE use DNA profiling. As such, it is necessary to say something about DNA evidence and why it is properly regarded as statistical.

Everyone knows that courts—particularly criminal courts—consider DNA evidence: but what is the nature of this evidence and how is it presented to the court? DNA evidence comes from extracting, analysing, and making statistical inferences from genetic material. A typical case is where such material (e.g. saliva, hair, semen) is found at a crime-scene—for instance, investigators finding a cigarette-butt containing saliva at the site of a break-in. This genetic material is processed in a laboratory to build a profile of which alleles—the genes we inherit from our parents—appear at which locations on a strand of DNA. From this process, DNA evidence is used in different contexts. One is where suspicion has already fallen upon some person and DNA profiling is used to bolster the case against them. This might involve taking further genetic material from a suspect—e.g. with a saliva swab—and testing it against the initial genetic material found at a crime scene. If the samples match, this further supports the pre-existing suspicion. A second use of genetic material, the

<sup>13</sup> See Enoch et al (2012: 198). Moreover, Blome-Tillmann (2017) suggests that such evidence should be inadmissible in defending what he views as a ‘conservative’ position on evidence law.

<sup>14</sup> For instance, see Smith (2018), Gardiner (2018), Di Bello (2019) for a selection of those who claim that using PSE would be legally revisionary.

use that we will primarily be concerned with, is so-called ‘cold-hit’ DNA evidence. In such cases genetic material (e.g. taken from the scene of a crime) is found to bear similarity to a DNA profile already held on file for some other reason. This evidence is then used as the basis of a criminal case against the person to whom the matching DNA profile belongs. In a cold-hit case, this person is *not* someone upon whom suspicion has independently fallen.

DNA evidence is presented to the court in the form of *expert testimony*. An expert witness, often a forensic scientist with a doctoral degree, is called to court to testify to the significance of the DNA evidence. In a cold-hit case, the most important element of the testimony concerns an estimated statistical frequency of the allele-configuration in a given target population: for example, here is Dr Clara O’Sullivan providing the key piece of inculpatory evidence in a recent high profile cold-hit DNA case:

**Expert Witness:** “We give an estimation of the chance of somebody randomly unrelated in the population, having the same profile, given that [the suspect has] the profile, is one in a thousand million.”  
[*Wilson v DPP* 2017 IESC 54 at 5.18]

In this sense, DNA evidence is statistical evidence: on the assumption that the genetic material taken from the crime-scene belongs to the person who committed the crime, the *incriminating component* of the DNA evidence is that it is exceptionally unlikely that the genetic material does not belong to the person being accused of the crime.<sup>15,16</sup> When there is no additional evidence apart from the statistical estimate based on the DNA profile—as was the case in *Wilson*—we have a case involving *purely* statistical evidence. With this in mind, we can return to consider how the law treats PSE.

#### 4.2. Purely statistical evidence in the law

There has been some confusion in the philosophical literature regarding statistical evidence in the law. One mistaken suggestion is that statistical evidence is generally *inadmissible*. If it were, we wouldn’t ever see criminal convictions relying on DNA evidence. Indeed, if statistical evidence were inadmissible, then it would be treated like other types of inadmissible evidence—such as evidence obtained through torture—i.e. it would not be found even in conjunction with non-statistical evidence.<sup>17</sup> This is clearly not the case: courts take statistical evidence into account all the time, on a range of different subject-matters. Of course, there might be particular cases in which certain types of statistical evidence *are* inadmissible—such as evidence pertaining to demographic crime-statistics—but this does not mean that statistical evidence is generally inadmissible.<sup>18</sup> These

<sup>15</sup> It is worth pointing out that even critics of PSE do not tend to argue that DNA evidence is non-statistical.

<sup>16</sup> To be explicit, the probability of error being testified to by the expert witness concerns the possibility of there being a *random match* in the target population who shares the relevant DNA characteristics as the sample, but was not involved in the relevant incident. Such estimates do not include other possibilities of error like the DNA sample being corrupted due to cross-contamination or laboratory mistakes.

<sup>17</sup> See *R. v Alan James Doheny; R. v Gary Adams (together)* [1997] 1 Cr. App. R. 369 for general rules on admitting DNA evidence.

<sup>18</sup> For instance, some cite *United States v Shonubi* 103, F.3d 1085 (2d Cir. 1997) in support of the inadmissibility of statistical evidence. However, this case clearly only concerns a very specific practice: punishing someone by using an estimate of contraband they have smuggled over time via a statistical evaluation of the habits of other smugglers. It does not follow that appellate courts disapprove of PSE in general just because they view such speculative sentencing as iniquitous.

points are neatly exemplified by a quote from the very same case we drew the example of expert testimony from earlier. The judges in this case, drawing on a range of common-law authorities from different legal systems, stressed the following about Dr O’Sullivan’s statistical testimony:

[I]t is appropriate to emphasise that the Court is not here concerned with the admissibility of such evidence. Clearly even evidence of a tangential connection may be relevant in the overall context of a particular case but would be unlikely to provide sufficient evidence, without more, to allow for a safe conviction. [*Ibid.* at 5.5]

As this quote suggests, the salient issue is whether PSE can discharge the evidential burden of providing *legally sufficient evidence*. Evidence can be admissible without being legally sufficient. A prosecutor can lead evidence that the accused owns a set of kitchen knives. Whilst admissible, this evidence would not be sufficient by itself to support a murder conviction.

So, have courts typically held that PSE can discharge the evidential burden? In *Wilson*, the judges answered in the affirmative. Indeed, if we focus on PSE in the form of criminal cases relying on DNA evidence alone, then there are now numerous examples of courts affirming the sufficiency of such evidence to undergird a conviction.<sup>19</sup> This is not to say that the sufficiency of PSE is an entirely uncontroversial matter.<sup>20</sup> However, it is important not to overplay the extent of the controversy. One mistake to avoid is citing *any* case in which a verdict based on purely statistical evidence is overturned as supporting the claim that legal practice *generally* eschews PSE. For instance, *R v. Watters*<sup>21</sup> is often cited as exemplifying the reluctance of the courts to rely on bare statistics. However, if one attends to the details of the judgement, the problem was not with statistical evidence *per se*, but rather with the fact that it failed to rule out the accused’s brother being responsible—and this was a possibility that had been raised by the defence. Indeed, the judges in *Watters* (at 21 per LJ Kay) are at pains to emphasise “we are not for one moment saying that merely because there was no other evidence . . . that this appeal has to be allowed”.

It is true that one can find cases where the courts express hostility towards PSE, especially when we move away from DNA profiling. The case of *Virginia & S.W. Ry. Co. v Hawk* in the civil domain—involving statistics about railway accidents—is a good example. However, other cases such as *Kaminsky v Hertz Corp*—superficially very similar to

<sup>19</sup> For a selection, see: *R v. Hanratty* [2002] 2 Cr. App. R. 30 (esp. comments at 127); *R v Adams (no1)* [1996] 2 Cr. App. R. 467 and *R v Adams (no2)* [1998] 1 Cr. App. R. 377; *R v Weir* [2000] 5 WLUK 751; *State v Toomes* 191 S.W.3d 122, 129 (Tenn. Crim. App. 2005); *State v Hunter* 861 N.E.2d 898 901; *State v Davis* 698 N.W.2d 823, 826-27. Also see *Director of Public Prosecutions v. Connolly* [2011] IESC 6 (which, although dismissed for other reasons, affirms the sufficiency of PSE in the form of a forensic scientist analysing packages containing amphetamine). Beyond these citations, Roth (2010) is a legal scholar who discusses a variety of cases in more detail than I have space for here. Further, in *Toomes* the judges cite a number of other historical cases in support of the sufficiency of statistical evidence, for the interested reader. It is also worth noting that many uses of cold-hit DNA leading to conviction go unreported, in particular cases in which the suspect pleads guilty.

<sup>20</sup> *R v. Lashley* [2000] EWCA 88 is an example of a DNA case being rejected on the basis that five or six people in the UK might fit the DNA profile. This doesn’t support a general rule against the sufficiency of DNA evidence however: often the probabilities involved in DNA cases make error even less likely than in *Lashley*.

<sup>21</sup> [2000 WL 1791491]

the BLUE BUS case as it involved ascribing liability on the basis of the percentage of vehicles owned by a particular company—where judges seem entirely amenable to purely statistical evidence. My own view is that such cases are not aspiring to be binding precedents about the standing of PSE in general, to hold for each and every case. Rather, they attend to specific details of the case at hand in making their judgement.<sup>22</sup> This is often how the common law develops: it proceeds in a piecemeal fashion by drawing fine distinctions between preceding cases as justice demands in the case at hand. However, getting into the details of legal interpretation is unnecessary to make the following broader point about the sufficiency of PSE: as the rise of DNA profiling shows, PSE is not insufficient evidence *in general* even though there may be individual cases in which the courts express reluctance to rely on it.

I have attempted to set the record straight on how PSE is treated in the law. This serves a number of purposes. Beyond the simple motivation of striving for accuracy, showing that PSE *is* currently used in the law rebuts the thought that there is universal distaste for PSE that philosophers merely have to vindicate. Rather, especially given the prevalence of DNA evidence, the proper status of PSE is a live question that we should consider with an open mind. Moreover, our discussion also allows us to sketch a sympathetic proposal on behalf of critics as to how their critique should be implemented. We should assume that critics of PSE are ill-served by suggesting that statistical evidence should be *inadmissible*. This would rule out using DNA evidence in any context, whilst also excluding statistical evidence when lead in conjunction with other types of evidence. I take it that this position is unappealing—there is little reason to suppose that statistical evidence has *no role* to play in legal decision-making. Rather, it is much more promising to interpret critics of PSE as arguing that our best evidence law should hold that: *statistical evidence alone cannot constitute legally sufficient evidence* to undergird a criminal conviction or finding of civil liability.<sup>23</sup> Under this approach, if the party with the burden of proof adduced only statistical evidence then the fact-finder (such as the jury) would not be asked to deliberate—rather, the case would be thrown out. This seems like the most plausible interpretation of how critics could have their proposal implemented by common law legal systems. Of course, as we have shown, such a proposal would be rather revisionary: in particular, it would disallow criminal convictions based solely on DNA evidence.

## 5. Doxastic approaches

The foregoing empirical facts about legal practice still leave the normative question open: should there be a blanket prohibition on bare statistics carrying the day in court? A number of philosophers provide an affirmative answer, seeking to identify features shared by all PSE cases that justify rejecting them. This section argues against three different approaches in this vein.

<sup>22</sup> For instance, it seems a better interpretation that *Virginia* concerns how juries should be *directed* in relation to probabilities: particularly, they should not be left to make probabilistic conjectures. *Kaminsky*, on the other hand, places weight on defeasible presumption created by the appearance of the truck and the unique demands of justice applicable to the law surrounding motor vehicles. Neither is a general endorsement or prohibition on PSE.

<sup>23</sup> Given that many juries (in the cases cited previously) have found PSE to satisfy the relevant SoP, it is also unappealing to suggest that it is a conceptual truth about any particular standard that it cannot be satisfied by statistics alone.

### 5.1. Three arguments against purely statistical evidence

A prominent line of thought—recently endorsed by Smith 2018; Littlejohn 2018; and Blome-Tillmann 2017 proposes to explain and vindicate the intuitions against PSE by appealing to the epistemic norms that apply to individual inquirers.<sup>24</sup>

Before looking at the specifics, we'll need some brief background theory. In epistemology there is an ongoing debate about how *individuals* should treat purely statistical evidence, best exemplified by preoccupation with lottery cases. Many suggest that there is an important normative asymmetry between statistical evidence and other types of evidence (such as direct sensory apprehension). For instance, some theories claim that relying on statistical evidence doesn't yield *knowledge*. Taking a concrete example, a dominant thought has been that no matter the size of a fair lottery, one cannot *know* that one has a losing ticket just on the basis of it being very likely.<sup>25</sup> Other theorists suggest that basing beliefs on purely statistical evidence leads to the absence of other normative properties, such as certain types of *justification*.<sup>26</sup> This concern with statistical evidence is one front in a broader dispute about which normative property should be central to the epistemic lives of individual inquirers. For instance, some endorse the idea that statistical evidence does not yield knowledge apiece with a more general view according to which knowledge determines when it is appropriate for someone to act upon a belief, to make an assertion, or even to hold a belief in the first place.

A number of philosophers have used ideas from epistemic theories concerning individual inquirers to argue against relying on purely statistical evidence in the law.

#### **Blome-Tillmann's argument**

Firstly, Michael Blome-Tillmann appeals to the idea that knowledge is the norm of action.

Assuming that we ought to act on  $p$  only if we know  $p$ , the ideal and strongest epistemic state a court could be in, from a normative point of view, is one in which the court knows the defendant is at fault. [Blome-Tillmann 2017: 284]

However, accepting that requiring knowledge is an unreasonable legal standard, Blome-Tillmann endorses an alternative view on which the *probability of possessing knowledge* is what matters:

Courts need to be able to make decisions in the absence of the epistemically normative ideal—that is, in the absence of knowledge that the defendant is at fault. What is needed is a measure of gradual approximation to the normative ideal. One way to measure such approximation is in terms of evidential probabilities that one knows that the defendant is at fault. Knowledge then remains the normatively ideal state, but epistemic success in courts of law can be measured and understood in terms of something less than knowledge—namely, the evidential probability that knowledge has been achieved. [Ibid.]

<sup>24</sup> These three support this view most explicitly. But it is also at least suggested by Enoch et al (2012); Bolinger (2018); Pritchard (2015); and Buchak (2014).

<sup>25</sup> Most prominently, see Williamson (2000). This thought has also been supported by empirical survey data (see Ebert et al. 2018). See Hawthorne (2003) for discussion.

<sup>26</sup> See Sutton (2007), Smithies (2002), or Smith (2016); each focus on lottery cases.

On this basis, Blome-Tillmann suggests that there should be a rule on which the SoP can only be satisfied by evidence that raises the probability that the conclusion is *known* above a certain threshold (e.g.  $>.5$  for the civil BoP standard). As, *ex hypothesi*, statistical evidence by itself cannot yield knowledge, PSE cases will not meet this threshold.

### *Smith's argument*

Martin Smith's argument requires a brief theoretical preface.

In his work on individual inquirers, Smith (2016) developed a concept of justification featuring a notion he calls *normic support*. Normic support can be introduced with the idea of something 'calling for explanation'.<sup>27</sup> According to this theory, you are normically justified in holding some belief  $p$  only if the falsity of  $p$ , given your evidence, would call for some special explanation. So, while it might be very unlikely that I have won the lottery based on probability alone, it wouldn't be *abnormal* for me to win in the sense of calling for some special explanation; thus, my belief that I have lost would not be normically supported. On the other hand, were I to *look* at my ticket and see that the numbers don't match the winning numbers, then my belief that I have lost would be normically supported: some special explanation would be called for, were I to be mistaken.

Smith observes that the puzzle of statistical evidence in the law can be explained by noting that the body of evidence involved fails to *normically support* the fault of the defendant. He then derives a normative claim from that observation:

[O]ur judgments about the presence or absence of normic support track our judgments about whether an affirmative legal verdict would be acceptable or unacceptable. What I suggest is a standard of proof that is met only if a proposition is normically supported by the evidence – only if the evidence makes the falsity of that proposition less normal, in the sense of calling for more explanation, than its truth. What I suggest is that a verdict of guilt or liability is only acceptable in so far as this normic standard is met. This is my proposed solution to the legal puzzle of statistical evidence. [Smith 2018: 1209-1210]

### *Littlejohn's argument*

Littlejohn holds that the reasonableness of a belief should be judged against the standard of acquiring knowledge, something that purely statistical evidence cannot yield. With this in mind, he endorses the following principle:

Reasonable Conviction: It is not permissible to punish a defendant if it isn't reasonable to believe the defendant to be guilty. [Littlejohn 2018: 15]

Littlejohn defends this principle primarily by appealing to considerations of legitimate punishment:

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. 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 considerations to be a punishment. [. . .] If we have a system of rules that governs decisions to punish or

<sup>27</sup> The formal version of Smith's theory appeals to a world-ranking framework, details of which can be found in his (2016).

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. [Ibid 16-17]

Littlejohn (s.8) expands on this suggestion by claiming that there are reactive attitudes and emotions associated with blame which can be rationalised by outright belief but not by a high degree of confidence that the subject of blame has transgressed.

### 5.2. Against doxastic approaches

Each of these arguments takes its cue from individual epistemology. Blome-Tillmann transposes to the law a theory about when it is rational for an individual to act. Smith applies to the legal domain his notion of epistemic justification which concerns when it is rational for an individual to believe something. And Littlejohn stakes out a position on when legal punishment is legitimate by appealing to the normative role of outright belief in licensing an individual to hold reactive attitudes associated with blame.

For what follows, I will call views fitting this mould *doxastic approaches* because they take their cue from the norms surrounding belief in individual inquirers. Doxastic approaches are favourable to **Never**. This is because, on doxastic views, relying on purely statistical evidence can never fulfil certain requirements: such as providing knowledge, rationalising action, legitimising blame, and/or providing normic support for some proposition. As such, their arguments do not rest on details of particular cases but rather appeal to generic facts about the epistemic power of purely statistical evidence. Given that the three arguments outlined each claim that courts should conform to the same normative standards as individuals when confronted with PSE, one way to object to them would take issue with how they characterise the norms governing individuals. I won't take this approach here. Rather, let's grant whichever epistemic view critics prefer; my argument will attack the purported normative symmetry between doxastic and legal norms.

Let's start with an observation: *in an ideal scenario*, when a court issues a judgement—e.g. that Harry killed Sally, or that Jim's hedge is encroaching Jules' property—it would be one that an individual inquirer could cheerfully believe outright on the evidence. This is because, in a perfect world, the evidence would always support the legal verdict without leaving any room for doubt or uncertainty. Thus, courts would only ever judge—and, act, blame and punish—when the evidence was strong enough to support an individual inquirer believing outright (or knowing) that the relevant party is at fault or not at fault. We do not exist in this ideal world. The reality is that courts are routinely faced with bodies of evidence that are difficult to evaluate and fail to unequivocally vindicate either side. This is why, for instance, courts cannot simply ask the fact-finder to look at the evidence and then report *whether they believe outright* that (i) the defending party is at fault, or (ii) the defending party is not at fault.<sup>28</sup> Such an approach would be a recipe for paralysis. Take a civil case where the evidence only weakly suggests that a

<sup>28</sup> To be clear: the doxastic views I critique below accept that such an approach is untenable. For instance, Littlejohn does *not* endorse a symmetrical view on which one must either know that a party is guilty *or* know that it is not guilty before judging accordingly—rather, his idea is that we should know the party is guilty before judging them guilty, and should exonerate otherwise. Thanks to anonymous referees for pressing me to clarify.

given party has been negligent; this will not be enough to support an individual believing outright that the relevant party either is, or is not, civilly liable. Nonetheless, it is incumbent on a tribunal to decide one way or the other. This is why legal tribunals have rules to ensure that they operate in a thoroughly binary fashion—something is treated as having occurred or not, an accused is judged guilty or not guilty, a defending party is deemed civilly liable or they are not.

Lord Hoffman, in an influential case on proof, puts it succinctly:

If a legal rule requires a fact to be proved (a “fact in issue”), a judge or jury must decide whether or not it happened. *There is no room for a finding that it might have happened.* The law operates a binary system in which the only values are 0 and 1. The fact either happened or it did not. . . . If the party who bears the burden of proof fails to discharge it, a value of 0 is returned and the fact is treated as not having happened. If he does discharge it, a value of 1 is returned and the fact is treated as having happened. [Lord Hoffman in *Re B* 2008 UKHL 35]

When the law treats something as having happened or not, this does not entail that the evidence supports an outright belief that the thing occurred or didn't occur. Even if the evidence is muddy, a final decision must be made to settle the case at hand. It is this *burden of deciding* that explains why evidence law contains standards and burdens of proof. As such, judges and juries do not need to form outright beliefs in order to issue verdicts: they need only look at the evidence lead by the party with the burden of proof and measure it against the relevant standard of proof (e.g. Jules might have the burden of showing, on the balance of probabilities, that he owns the land that Jim's hedge is growing on). It is worth underlining this point because there is an assumption explicit in some defences of doxastic approaches (e.g. when Blome-Tillmann writes about a court knowing that the defendant is at fault) and implicit in others (e.g. when Littlejohn discusses the role of beliefs in rationalising punishment) that legal verdicts are a species—or at least close analogue—of belief. However, legal judgements and beliefs are crucially dissimilar: whilst belief is often thought to be a *two*-place relation between an individual and some proposition, legal judgements involve a *three*-place relation between an *agent* (the fact-finder: whether in the guise of jury or judge), some *proposition* (e.g. the civil liability of a defendant), and a *standard of proof* (e.g. the balance of probabilities). The fact-finder in a court is not asking whether they believe a proposition simpliciter, but rather deciding whether the evidence supports that proposition on a given evidential standard. Meeting this standard *may or may not* mean that she believes the proposition outright.

Arguments against purely statistical evidence can nonetheless be reframed entirely in terms of *evidential support*. On this approach, critics can say that legal verdicts must be based on evidence that could yield knowledge or normically supported belief in an individual. However, this position begs the following question: why should we suppose that norms regarding what individuals should do on the basis of certain types of evidence will automatically transfer to legal tribunals? After all, it seems reasonable to suppose that the conditions under which an individual should judge or act might differ from when a legal tribunal should judge or act. Of course, both courts and individuals will be governed by *some* norms of epistemic rationality. But we should question whether these norms are *identical*. For, if there turn out to be significant differences between the epistemic predicaments of individuals and legal tribunals, then extant arguments against legal reliance on bare statistics will miss the mark.

Recall the fact that legal verdicts are binary. In light of this fact, when it comes to the question of how to act—and whether to judge someone responsible for something as a preface to blaming or sanctioning them—individuals and courts face different options. One crucial difference is that a legal tribunal must decide one way or the other, whilst individuals are able to *hedge*. Let's illustrate with an example.

*Neighbour.* Suppose that you come across a confidential file stating that there is DNA evidence linking your new neighbour to a series of sadistic and violent murders. The file contains expert testimony stating that DNA material supports a 999,999/1,000,000 chance of him having committed these awful crimes.

Should you believe outright that he is guilty? Well, from the perspective of addressing various practical questions, there is a sense in which it is unimportant. After acquiring this evidence, will you ask him to babysit your children, to look after your spare key, or perhaps accept his offer to go hiking—just the two of you—in the backcountry? Certainly not. As an individual with statistical evidence indicating the stupendous likelihood that your neighbour is a serial killer, you will adopt various prudential policies, rejecting his offer to share a late-night walk through the moors. This is true even on the assumption that purely statistical evidence doesn't license outright/normically supported/knowledgeable belief. Even if you only believe <he is extremely likely to be guilty> rather than believe that he is guilty outright, you will change your behaviour considerably.<sup>29</sup> Moreover, this response is perfectly *rational*—indeed, your partner may rightly criticise you if, for example, you give the neighbour the spare key for safe-keeping. This is true, even if everyone concedes, “Of course, we don't *know* that he's a killer!” The ability to hedge is very important: to further illustrate, consider that many norms relied on by critics of statistical evidence were developed to deal with lottery-cases. However, individuals faced with deciding whether or not they have won the lottery are not compelled to return a binary verdict of <yes> or <no>. We are not forced to *either* act as if we have lost and throw away the ticket, *or* act as if we have won and sign up for early retirement. Rather, an individual can maintain a hedged belief that they have *probably* lost—thus rationally licensing them to hold on to the ticket but not to rely upon potential winnings in making financial plans.

Legal tribunals, faced with statistical evidence, are in a rather different epistemic predicament. Courts cannot judge that a party is stupendously likely to be guilty: rather, they must issue a judgement of guilt or innocence.<sup>30</sup> A legal tribunal cannot forbear from ascribing fault *and* act in a way that reflects a high degree of confidence in fault. So, we are owed an argument as to why the epistemic norms governing individuals (who *can* hedge) should result in a legal tribunal (which *cannot* hedge) exonerating a suspect when the evidence is purely statistical. For, even if it is true that an individual should believe and act on the proposition <probably *p*> rather than <*p*> upon considering bare statistics supporting *p*, this is not an effective argument for supposing that a court should issue and act on the judgement <not-*p*> on the basis of the same evidence. After all, we *do*

<sup>29</sup> I put my point in terms of probabilistic belief, but it works equally well in a credence-based framework.

<sup>30</sup> A fascinating exception here is the *third verdict* in Scottish criminal cases of 'not proven'. The existence of this verdict is deeply controversial from the perspective of criminal law theory, and its abolition has been mooted at various points. I won't take any stance here. In any case, the existence of this third verdict in one legal system doesn't cause trouble for my general point—a 'not proven' verdict is functionally identical to a 'not guilty' verdict insofar as it doesn't license any additional action by the legal system.

think it is rational for individuals to treat people differently after gaining statistical evidence about their possible conduct. Legal tribunals cannot navigate between extremes by hedging like individuals can, because courts have the burden of issuing a decisive verdict. The normative position of a tribunal facing statistical evidence is thoroughly different from that of an individual—hence, if you are attempting to explain why a court should not find someone guilty on the basis of bare statistics, it is not enough to simply appeal to how things stand with individuals.

## 6. Rehabilitating statistical evidence

Let's now make progress on what courts *should* do in response to purely statistical evidence.

Although courts must issue a binary verdict, the intuitions behind the puzzle of PSE suggest that it isn't *obviously* appropriate to simply decide in favour of whatever conclusion has most probabilistic support on the evidence. Here's a first stab at explaining why this is: courts have a very special role in doling out criminal and civil penalties, and this role brings with it a host of competing responsibilities. Consider a criminal court. If it acts on a high probability that someone is guilty (e.g. by punishing them) when they are in fact innocent, then it has wronged that party. Moreover, given the public nature of legal judgements, false convictions threaten the credibility of the legal system itself.<sup>31</sup> Courts must aim to both avoid wronging those they punish and to uphold confidence in the legal system. However, legal systems also have duties that can be frustrated by exonerating the accused. If it exonerates a guilty party, then the court has failed in its aim of delivering the appropriate response to crime (whether fleshed out in terms of protecting society from further criminal activity, deterring crime, giving due retribution, providing rehabilitative opportunity, or performing some semiotic function). Hence, courts facing purely statistical evidence are in the epistemic predicament of being obliged to issue a binary decision where there are unfortunate consequences of getting it wrong in either direction. The norms of legal inquiry are encumbered with the delicate task of managing these risks.

Dealing with the possibility of error in a satisfactory way within the binary framework constraining tribunals is a jurisprudential question facing every legal system. In this sense, legal systems cannot only be concerned with accuracy. When setting the relevant evidential standards they must also take a stance on what type of errors should be minimised at the expense of others. Consider the BRD standard of proof. The evidential standards required for criminal conviction must strike a difficult balance. First and foremost, legal systems need to avoid falsely convicting the innocent. Indeed, in criminal law theory, it is typically thought to be worse to falsely convict the innocent than mistakenly acquit the guilty. However, avoiding false convictions is not the only imperative for the criminal law: otherwise we would endorse a beyond *any* doubt standard of proof. Clearly, it is essential to ensure that those who are guilty of crimes have an acceptable chance of being convicted. The BRD standard is supposed to strike a balance between the desiderata of avoiding false convictions whilst still ensuring that conviction for criminal activity is a realistic prospect. That the BRD standard strikes the best balance is far

<sup>31</sup> Indeed, some (e.g. Thomson 1986; Bolinger 2018) point out that legal verdicts resemble *assertions*. As such, we might expect their status as assertions to generate additional norms. I won't rely on this thought here, but this is one way to precisify my argument if you are tempted by the verdict/assertion parallel.

from a platitude. For example, Laudan (2006) has suggested—based on the idea that one is more likely to be a victim of a recidivist criminal than falsely convicted—that the standard should be weakened. Regardless of where the boundary is set, some wrongful convictions and false acquittals are inevitable; the question is how to apportion these risks.

With this in mind, I want to suggest the following about cases involving purely statistical evidence: our intuitive discomfort about these cases stems from the fact that such evidence makes the possibility of error salient. This makes us reluctant to sanction, because any error would constitute an injustice against the sanctioned party. Nonetheless, it is part and parcel of recognising the different responsibilities that legal systems have, and the fallibility of legal tribunals, that we tolerate the fact that courts will sometimes make errors when discharging the burden of deciding one way or the other. Satisficing the different aims of the legal system requires accepting that we will sometimes convict people for crimes they did not commit. When confronted with cases that throw this fact into sharp relief it naturally makes us uncomfortable, because we are understandably hesitant to explicitly endorse trade-offs in which the innocent will foreseeably suffer miscarriages of justice to facilitate justice being done elsewhere. However, this balancing act is something that all legal systems *must* countenance in virtue of their unique role in settling disputes. In contrast, the possibility of error is usually much less salient when relying on non-statistical evidence. For, when we rely on evidence such as eye-witness testimony, it is much easier to construct a psychologically persuasive narrative that obscures the possibility that the verdict is mistaken.

My contention about error-salience chimes with both philosophical speculation and empirical evidence. For instance, John Hawthorne (2003: 15-20) suggests that certain scenarios involving statistical evidence are structured in such a way as to engender what he calls *parity reasoning*. Parity reasoning occurs when we consider a proposition  $p$  against the thought that some particular  $p_i$  in a set of subcases  $p_1, \dots, p_n$  might not obtain and then realise that our evidence doesn't speak in favour of any particular instance not obtaining over any other particular instance. (Imagine this with respect to lottery-tickets: the evidence doesn't privilege any ticket—or 'subcase'—not losing over any other.) The result of parity reasoning is to make the chance of getting things wrong salient; after all, the evidence doesn't privilege the conclusion that any *particular*  $p_i$  in the set will not obtain. While ordinary non-statistical evidence doesn't lend itself to being conceptualised in this structured way, lottery cases clearly inspire parity reasoning because we naturally just compare our generic ticket to the other generic tickets, each with the same chance of winning. I think something similar occurs when there is inculpatory statistical evidence. Consider a cold-hit DNA case. When we hear expert testimony asserting "There is a 1/1,000,000 chance of this DNA belonging to someone else", I think we are moved to compare our current epistemic position with how things would be in other close possible worlds (or subcases), including one in which the person was innocent. By making this comparison, we realise that our evidence doesn't privilege any particular subcase and hence the possibility of error (i.e. a false conviction) is made very salient. This thought about salience is also supported by work from psychologists who have studied the Wells' effect: for instance, Neidermeier et al. (1999) suggest that juror reluctance regarding PSE is explained by the ease of imagining a scenario in which the defendant is not at fault. This is all consistent with the empirically well-supported 'story model' of decision-making—a popular paradigm for explaining how juries make decisions in trials—on which legal verdicts are made through imposing a linear narrative onto the evidence presented

at trial.<sup>32</sup> Judgements based on bare statistics are not easily subsumed into a psychologically compelling narrative, and PSE does not lend itself to inclusion in a story more plausible than an opposing story in which the accused was simply absent.

If something like the salience thought is right, what we should do about our intuitive discomfort with PSE? One reaction is to simply disregard our intuitions, favouring PSE because it could bring about greater accuracy in the legal system overall. However, I do not think my argument compels us to take this position. This is because legal systems prize virtues other than accuracy, as is richly exemplified by different canons of evidence law. For example, some rules on evidential admissibility serve broader goals not closely tied to maximising accuracy. Consider the exclusion of evidence gained through impropriety—what US law calls ‘the fruit of the poisoned tree’—such as where some permission (e.g. a search warrant) was needed but not obtained. This evidence is excluded, even if evidence acquired through an inappropriate search might be highly reliable in a given case, or indeed highly reliable in general. Such rules serve aims such as discouraging inappropriate police conduct but they also serve the broader aim of ensuring *public confidence* in the legal system—an imperative embodied in the familiar maxim that ‘justice must not only be done, but be seen to be done’. Evidence law is shot through with rules that are not about maximising local accuracy.<sup>33</sup>

Let’s now return to the PRISONERS case.

PRISONERS: 100 prisoners are exercising in the prison yard. 99 of them attack the guard, putting into action a plan that the 100<sup>th</sup> prisoner knew nothing about. The 100<sup>th</sup> 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.

One argument *against* convicting in PRISONERS that is consistent with all we have said so far is the following: guilty verdicts in these types of case would offend certain desiderata of evidence law other than maximising accuracy. An obvious candidate is the imperative of maintaining public confidence in the legal system. In this sense, simply the fact that there is a *salient perception of possible injustice* in the PRISONERS case could be good enough reason to hold that the evidence is insufficient to discharge the evidential burden. This line of thought can be bolstered by observing that the sacrifices required by consistently exonerating in cases like PRISONERS are not egregiously costly: cases like PRISONERS are relatively rare, excluding such cases would not predictably enable people to escape justice for certain types of recurring crime, and—assuming that ‘reasonable doubt’ and a 1/100 likelihood of innocence aren’t *radically* divergent standards<sup>34</sup>—such a rule would not substantially alter the global accuracy of the legal system.

However, the same considerations do not apply across the board. In particular, even if we suppose that we can vindicate exoneration in PRISONERS on non-accuracy based grounds, I think we should resist this thought regarding (at least some applications of) DNA evidence. Therefore, my argument is one in favour of **Sometimes** rather than

<sup>32</sup> See Pennington and Hastie (1994) for an overview and discussion of experimental results.

<sup>33</sup> For other examples, consider rules precluding spouses being compelled to testify against each other; evidence gathered through police coercion; and so forth.

<sup>34</sup> For instance, surveys of judges on their view of the BRD standard (discussed in Solan 1999) support this assumption.

**Never.** Consider the following case fitting the mould of prominent uses of PSE in recent years:

*DNA:* Someone is sexually assaulted in a secluded park. They are unable to provide an informative account of the appearance of the attacker. Their injuries show that there is no question that the act was non-consensual. DNA evidence from the crime matches that of someone on file for some other reason. A forensic scientist estimates the chance of the DNA *not* belonging to that person to be 1 in one billion.

I do not think that this type of case creates the same perception of possible injustice as the PRISONERS case. One reason for this—whether well-founded or not—is that people tend to have faith in the scientific credentials of DNA evidence. This might partly be due to a misconception that the inculpatory component of DNA evidence is not merely statistical. However, even if people are directed to focus on probabilities alone, there has been empirical evidence suggesting that people are more inclined to think that legal verdicts based on PSE are appropriate the higher the probabilities of accuracy are (see Wright et al. 1996).<sup>35,36</sup> In addition, the sacrifices entailed by consistently exonerating in such DNA cases are more costly than a similar policy regarding PRISONERS style scenarios. Firstly, cases—particularly sexual offences—where the only incriminating evidence is DNA material are not particularly rare. Secondly, the lack of non-DNA evidence is a particular problem in prosecuting certain types of sexual offence, viz. those carried out by strangers with no particular motive in areas where there are unlikely to be eye-witnesses. Ruling DNA evidence to be insufficient would predictably undermine our efforts to prosecute a particular type of case, one that already suffers from perennially low conviction rates. And finally, a rule against this PSE in the DNA case *would* upset the global balance of accuracy in the legal system: a 1 in one billion chance of innocence is not a remotely reasonable approximation of the notion of a reasonable doubt, thus such a rule would lead to many more false acquittals than we would expect from allowing courts to use such evidence. So, drawing on these different factors as justification, we can still endorse the sufficiency of DNA evidence to undergird criminal convictions whilst vindicating the negative intuition about PRISONERS.<sup>37</sup>

My discussion focused primarily on criminal cases, but broadly the same points apply to civil cases too. The law of evidence regulating civil cases also involves a delicate

<sup>35</sup> Ebert et al. (2018) found regarding lottery cases that subjects are more likely to ascribe justified belief (but not knowledge) when the lottery is much larger—i.e. when the chance of error is much lower. This doesn't directly support any hypothesis about legal cases. However, it does support the thought that PSE suggesting an *extremely low* chance of error will receive a different (and, indeed, a more positive) evaluation than PSE merely suggesting a low chance of error.

<sup>36</sup> Bear in mind that an accused can lead exculpatory evidence in their favour: particularly, evidence speaking to an alibi. Of course, one might worry that it would be difficult even for an innocent person to overcome such formidable inculpatory statistics. But this worry is not particular to DNA evidence, for it is not always the case that an innocent person can easily rebut incriminating evidence against them. For instance, if someone is the victim of a sustained and skilful attempt to frame them, then various types of evidence (manipulated CCTV, bribed witnesses, etc.) will not easily be undermined. This is a consequence of trials being fallible, not a reason to regard any class of typically reliable evidence as generally unsatisfactory.

<sup>37</sup> There is nothing sui generis about DNA evidence—rather, using PSE raises different issues in different contexts and each should be considered on its individual merits against the broader aims of the legal system.

balancing act. Even though criminal sanctions such as imprisonment or the failure to punish crime is not at stake, the issues raised are nonetheless of substantial import. Take BLUE BUS style cases involving a harm being suffered by one party (i.e. being seriously injured in a bus accident). Evidence law must strike a balance between ensuring that people can be compensated for various (often severe) harms that result from the activity of others, and ensuring that others are not easily held liable for these harms erroneously. To put this point in context: either the party who has been struck by a bus receives compensation (e.g. for loss of earnings and other expenses incurred as a result of injury) or they are left with nothing. My own thought about BLUE BUS-type cases is that bare statistics *could* be enough to legitimately result in a finding of liability—sometimes it might be reasonable to distribute economic risks in such a way as to fall on a company with a predominant market share. Of course, this suggestion is defeasible; e.g. we should not take such an approach if it would predictably lead to iniquity for one company over others. However, it doesn't seem *necessarily* unfair to ascribe civil liability on the basis of statistics simpliciter.<sup>38</sup> We can press this thought by conceiving of civil cases involving very long odds more akin to DNA cases. For instance, imagine a case where there is only one bus company operating in a given locale. If it were the case that, say, only 1/10,000 buses on the road is owned by a private individual, then it could be adduced as evidence that there is a 99.99% chance of the bus that caused an accident being owned by the MONOPOLY BUS COMPANY. Again, I think that a finding of liability here is entirely reasonable. However, accepting this contrarian position is not really needed to vindicate **Sometimes** in the civil domain. For, it follows from what I have already argued—namely that criminal convictions can sometimes be appropriately based on PSE—that the same holds in the civil domain too. This is because many actions that constitute criminal offences—e.g. sexual offences or physical assaults—are also civil wrongs. Hence, it is a platitude that if PSE can rightly undergird a criminal conviction on the BRD standard then it must also be sufficient to support a finding of civil liability on the weaker BoP standard.

## 7. Conclusion

This paper examined a pressing legal question, namely: how should the courts treat evidence that is purely statistical in nature? A growing number of philosophers have suggested that legal systems tend to scorn such evidence and provided overarching theories attempting to vindicate this purported truism about the law. Against these views, I demonstrated that legal systems do in fact rely on purely statistical evidence: DNA profiling is a prominent example. Then I argued that one guiding thought behind many critiques of purely statistical evidence, namely that courts should be subject to the same epistemic norms as individual inquirers, is unconvincing. This is because courts, unlike individuals, cannot *hedge* their decisions—they must settle matters one way or the other by issuing a decisive verdict. And finally, I argued that the best approach to purely statistical evidence takes each case on its merits and illustrated that we can sensibly eschew

---

<sup>38</sup> Recall *Kaminsky v Hertz* in which a Blue Bus style scenario led to a finding of liability. The case explicitly considers the unique demands of justice raised by transportation law. This underscores my points: appropriate distribution of risk is case-specific, and my piecemeal approach is in keeping with how the common law develops.

such evidence in the sorts of cases that critics find most uncomfortable, but vindicate this evidence in other contexts such as the use of DNA profiling.<sup>39</sup>

### Bibliography

- Basu, Rima (2019). The Wrongs of Racist Beliefs. *Philosophical Studies*:1–19.
- Blome-Tillmann, Michael (2015) ‘Sensitivity, Causality, and Statistical Evidence in Courts of Law’ *Thought* 4(2): 102–112.
- (2017). ‘More Likely Than Not’ - Knowledge First and the Role of Statistical Evidence in Courts of Law. In Adam Carter, Emma Gordon & Benjamin Jarvis (eds.), *Knowledge First - Approaches in Epistemology and Mind*. Oxford, UK: Oxford University Press. pp. 278–292.
- Bolinger, Renée Jorgensen (2018). The rational impermissibility of accepting racial generalizations. *Synthese*:1–17.
- Buchak, Lara (2014). Belief, credence, and norms. *Philosophical Studies* 169 (2):1–27.
- Di Bello, Marcello (2019). Trial by Statistics: Is a High Probability of Guilt Enough to Convict? *Mind*:fzy026.
- Ebert, Philip A., Martin Smith, and Ian Durbach (2018). ‘Lottery Judgments: A Philosophical and Experimental Study’. *Philosophical Psychology*, 31(1), pp. 110–38.
- Enoch, David, Spectre, Levi, & Fisher, Talia (2012). Statistical Evidence, Sensitivity, and the Legal Value of Knowledge. *Philosophy and Public Affairs* 40 (3):197–224.
- Gardiner, Georgi. (2018). ‘Legal Burdens of Proof and Statistical Evidence’. In David Coady & James Chase (eds.), *The Routledge Handbook of Applied Epistemology*. London: Routledge.
- Hawthorne, John (2003). *Knowledge and Lotteries*. Oxford University Press.
- Laudan, Larry (2006). *Truth, Error, and Criminal Law: An Essay in Legal Epistemology*. CUP
- Littlejohn, Clayton (2018). Truth, Knowledge, and the Standard of Proof in Criminal Law. *Synthese*:1–34.
- Mogensen, Andreas (2019). Racial Profiling And Cumulative Injustice. *Philosophy and Phenomenological Research*.
- Niedermeier, Keith, Norbert Kerr, & Lawrence Messe (1999). Jurors’ Use of Naked Statistical Evidence: Exploring Bases and Implications of the Wells Effect. *Journal of Personality and Social Psychology* Vol. 76, 4; 533–542
- Pennington and Hastie (1994) The story model for juror decision making. In Hastie, R (Ed.). (1994). *Inside the Juror: The Psychology of Juror Decision Making*. Cambridge: Cambridge University Press.
- Pritchard, Duncan (2015). Risk. *Metaphilosophy* 46 (3):436–461.
- Redmayne, Mike (2008). ‘Exploring the Proof Paradoxes’. *Legal Theory*, 14(4), pp. 281–309.

<sup>39</sup> Thanks, as always, to my friends and colleagues at the University of St. Andrews for discussion and support. An audience at the University of Edinburgh provided helpful comments on a penultimate draft. I’d also like to especially thank Marvin Backes, Hannah Rose Blakeley, Jessica Brown, and Martin Smith for help with this paper. Martin did an excellent job of introducing me to epistemology as an undergraduate at the University of Glasgow—it is very gratifying to now be in a position to engage with him in print.

- Roth, Andrea (2010). Safety in Numbers - Deciding When DNA Alone is Enough to Convict, 85 *NYU Law Review* 1130
- Smith, Martin (2016). *Between Probability and Certainty: What Justifies Belief*. Oxford University Press UK.
- (2018). ‘When Does Evidence Suffice for Conviction?’ *Mind*. <https://doi.org/10.1093/mind/fzx026>.
- Smithies, D. (2012). The normative role of knowledge. *Noûs*, 46, 265–288.
- Solan, Lawrence M. (1999). Refocusing the Burden of Proof in Criminal Cases: Some Doubt About Reasonable Doubt, 78 *Texas Law Review* 105, 126–27
- Sutton, J. (2007). *Without justification*. Cambridge, MA: MIT Press
- Thomson, Judith Jarvis (1986) ‘Liability and Individualized Evidence’ *Law and Contemporary Problems* 49(3): 199–219
- Wells, G.L. (1992). Naked statistical evidence of liability: Is subjective probability enough? *Journal of Personality and Social Psychology*, 62, 739–752.
- Williamson, Timothy (2000). *Knowledge and its Limits*. Oxford University Press.
- Wright, Edward, Lora Maceachern, Elaine Stoffer, & Nancy Macdonald (1996). Factors Affecting the Use of Naked Statistical Evidence of Liability, *The Journal of Social Psychology*. 136: 677–685