The Burden of Proof and Statistical Evidence

GEORGI GARDINER

Rutgers University

This essay is forthcoming in the *Routledge Handbook of Applied Epistemology*, eds. David Coady and James Chase. Routledge.

1. The Golden Thread

The presumption of innocence is a keystone of many criminal justice systems. In criminal trials, the defendant is formally considered 'innocent until proven guilty'. The burden of proof is on the state to prove the defendant committed all elements of the crime; the defendant, by contrast, need not prove his innocence. This asymmetry was famously dubbed 'the golden thread' of criminal law.[1] Correspondingly in many legal systems, such as Anglo-American common law, the defendant in civil cases enjoys a presumption of non-culpability; the burden falls to the plaintiff to establish the facts of the case.[2]

A legal burden of proof generates a formal standard of proof. This is the standard to which the facts must be established to satisfy the burden. Formal standards include the 'reasonable suspicion' standard in the US, which is required for a brief stop and search.[3] A slightly higher standard, 'probable cause', is required for detentions, arrests and indictments, and for more substantial searches of persons and property. When quantified, the thresholds for satisfying the standards are typically glossed as around 10–50% and 30–60% confidence, respectively, that the person is participating in criminal activity.[4]

The 'preponderance of evidence' standard, usually understood as requiring that a claim is more likely true than not, operates in civil and family courts. The standard is typically quantified as exceeding 50% likelihood that the litigated facts obtain.[5] The 'clear and convincing evidence' standard, frequently quantified as around 65–75% confidence, is more demanding. This standard is employed in equity cases such as right-to-die hearings, wills, libel, child custody, paternity disputes, and commitment to mental institutions.[6]

The proof 'beyond reasonable doubt' standard, which is required for criminal conviction in many jurisdictions, is more demanding still. It is one of the most familiar aspects of criminal

---

[1]   Lord Sankey in Woolmington v DPP (1935) AC 462.
[2]   The burden of proof for some claims – such as affirmative defences or immunities from prosecution – falls to the defendant.
[3]   Terry v. Ohio, 392 U.S. 1 (1968).
[4]   McCauliff (1982); Lerner (2003); Bacigal (2004); Goldberg (2013). Note there can be substantial difference between an officer's confidence that crime occurred and her confidence that a specific individual committed that crime.
[5]   This standard is also known as the 'balance of probabilities' standard (In re Winship, 397 U.S. at 371–72; House of Lords in *Re B (A Child)* (2008) UKHL 35). For information about quantifying the standard, see McCauliff (1982); Simon (1969).
[6]   See McCauliff (1982); Vars (2010); Simon and Mahan (1971); United States v. Fatico, 458 F Supp. 388 (E.D.N.Y. 1978); Sand and Rose (2003).

law and is commonly understood to be around 90–95% confidence in the guilt of the defendant.[7]

Burdens of proof and their associated standards raise many epistemological questions. Questions concern whether the presumption of innocence is a propositional attitude, the nature of this attitude or stance, and whether it is appropriate. There are questions about how to interpret, and what legitimates, the various standards of proof.[8] In this essay I examine whether these standards can be properly understood in quantitative terms and how this affects the kinds of evidence that can satisfy the standard. In section two I introduce examples that suggest legal standards of proof cannot be properly interpreted in quantitative terms. In sections three, four, and five I evaluate three competing proposals for a non-quantitative epistemic condition on satisfying legal burdens of proof.

2. Statistical Evidence
Consider the following cases:[9]

> Underage Alcohol. Reliable studies establish that at most undergraduate house parties, underage drinking is ubiquitous. The local police force decides this base rate evidence satisfies the reasonable suspicion standard. On this basis they search underage individuals arriving at undergraduate parties.[10]

> Gatecrasher. A music venue sells seats at its event but does not issue, or record who purchases, tickets. One day the gate is left open and, taking advantage of the lack of ticketing system, many people gatecrash. The managers realise only 10 seats were sold, but 80 people attended. They call the police, who arrest some attendees for gatecrashing. The police reason that there is probable cause to arrest any attendee, since it is overwhelmingly likely they committed a crime.[11]

> Red Taxi. A vehicle hit Jeanette late one night. She could determine it was a taxi, but could not discern the colour. The Red Taxi company operates 75% of taxis in town. The remaining 25% are operated by the Green Taxi company. Jeanette sues the Red Taxi company. Using only the evidence described here, Jeanette reasons, she will win the case, because it is more likely than not that Red Taxi is liable.[12]

---

[7] See, for instance, Tribe (1971); Laudan (2011); Walen (2015); Mulrine (1997); Lippke (2010).

[8] See, for example, Laudan (2011; 2012) and rebuttals by Risinger (2010) and Gardiner (2017a). See also Lerner (2003); Walen (2015); Mulrine (1997)l Nesson (1979); Pardo and Allen (2008); Sand and Rose (2003); and the 2006 exchanges in *Law, Probability, and Risk* (vol. 5), including Newman (2006), Franklin (2006), Tillers and Gottfried (2006), and Weinstein and Dewsbury (2006).

[9] I include five kinds of example partly because readers unconvinced by one might find a different example compelling, and partly to provide an example for each standard of proof outlined above. For related examples and discussion, see Tribe (1971); Redmayne (2008); Koehler (2001); Schneps and Colmez (2013); Thomson (1986); Buchak (2014); and People v. Collins 68 Cal.2d 319, 66 Cal.Rptr. 497 (1968).

[10] This case is inspired by Kerr (2012). To satisfy the reasonable suspicion standard the officer must have a specific, articulable, and individualised suspicion that crime is afoot before they conduct a brief search (Terry v. Ohio, 392 U.S. 1 (1968)). Police might reason the statistic-based suspicion is individualised: those particular underage individuals enter an undergraduate party.

[11] Cohen (1977). Note the example is usually employed to illuminate the preponderance of evidence standard rather than probable cause.

[12] This case was first discussed in print in Tribe (1971). Tribe refers to the case as a "famous chestnut" (p. 1341). See also Thomson (1986); Nesson (1985); Smith v. Rapid Transit 317 Mass. 469, 58 N.E.2d 754 (1945); Kaminsky v. Hertz Corp. 288 N.W. 2d 426 (Mich. Ct. App., 1980).

> Uncertain Paternity. Magda does not know who fathered her child. She knows it is one of two lovers, and asks them to undergo DNA testing. Owing to disorganisation at the laboratory, both samples become adulterated. Statistical analysis cannot generate the high certainty characteristic of flawless DNA testing but reports a 90% likelihood that Tom is the father. Since paternity tests are expensive, Magda reasons this evidence ought suffice for her paternity suit, which only requires 'clear and convincing evidence'.

> Prison Yard. One hundred prisoners exercise in the prison yard. Ninety-nine prisoners together initiate a premeditated attack on a guard. Security footage reveals one prisoner standing against the wall refusing to participate. There is no evidence indicating who refused to participate. The prison officials decide that since for each prisoner it is 99% likely they are guilty, they have adequate evidence to successfully prosecute individual prisoners for assault.[13]

In each case brute statistical evidence plausibly licenses a relatively high credence in the target claim. If the burden of proof can be satisfied by statistical evidence alone, these vignettes exemplify satisfying the relevant burdens. But there is something dubious about these cases. Law courts would not adjudicate in favour of the claimants and the police would likely violate their code of conduct were they to act on this basis.[14]

For each vignette, furthermore, we can construct a comparison case in which non-statistical, 'individualised' evidence is employed instead. The overall likelihood of the disputed fact given the evidence might be lower, yet the burden seems satisfied. Consider the following case:

> Red Taxi Testimony. A vehicle hit Jeanette late one night. The Red Taxi company operates 50% of taxis in town. The Green Taxi company operates the remaining 50%. Jeanette claims she saw the taxi was red. Eyewitness reports are notoriously unreliable, and so tests are performed to determine Jeanette's reliability under the relevant conditions. She discerns the correct colour 70% of the time. Her eyewitness testimony thus suggests a red taxi caused the accident. With this evidence Jeanette sues the Red Taxi company.[15]

In some sense, this eyewitness evidence is less determinative: plausibly a red taxi is more likely at fault given the mere market share evidence in the original example. But finding liability in the second case seems more legitimate and courts are significantly more likely to find in the plaintiff's favour. Similar counterparts can be constructed for each of the five vignettes above.

These cases suggest the standard of proof is not a quantitative measure, such as a credence or statistical likelihood. If 'preponderance of evidence' could be captured numerically, for example, it would be lower than 75%. And this would mean, implausibly, the fact finder should find liability in the Red Taxi case. Similarly if 'beyond reasonable doubt' can be understood quantitatively, it must be lower than 99% confidence. This would entail, implausibly, statistical evidence in the prison yard case suffices for conviction. Questions about whether merely statistical evidence can satisfy a burden of proof are increasingly pressing, as cold hit DNA or fingerprint matches and other statistical approaches to

---

13   Nesson (1979).
14   Koehler (2001). See also Gardiner (forthcoming), which casts doubt on the adequacy of merely statistical evidence in non-legal contexts.
15   Thomson (1986). For related discussion see Tversky and Kahneman (1982).

evidence are becoming more prevalent.[16] ('Cold hits' are when a database search connects a hitherto unconnected perpetrator to a crime.) This raises the question: if the standard of proof is not simply a numerical likelihood, and if merely statistical evidence does not satisfy that standard, what other conditions are required to satisfy the burden of proof?[17]

3. Causal Relations and Guarantees

Legal theorists note that vignettes like the first five above lack 'individualised' evidence against the defendant. But this notion is itself opaque. Judith Jarvis Thomson aims to illuminate the nature and value of individualised evidence.[18] She argues individualised evidence is that which is appropriately causally connected to the target claim. This causal relation might be that the evidence caused the crime. Gambling debt, for example, might be individualised evidence of guilt if debt motivated the theft. Thomson calls this 'forward-looking' individualised evidence. Alternatively crime can cause evidence. The crime's occurring, for example, causes incriminating fingerprints. Thomson calls this 'backward-looking' individualised evidence. Finally, evidence and crime might have a common cause. Suppose the defendant verbally abused an officer one evening. This might be evidence the defendant later attacked a pedestrian because it indicates a common cause – the defendant was enraged that evening – for both incidents.

Evidence with an appropriate causal relation, Thomson argues, supplies epistemic import distinct from merely raising the likelihood of the disputed claim. Individualised evidence, in her view, 'guarantees' the claim is true.[19] In the Red Taxi Testimony case, the taxi's being red causally explains the eyewitness evidence. Eyewitness evidence can therefore guarantee, argues Thomson, the litigated taxi-colour claim. Thomson writes, 'to require individualized evidence of guilt just is to be requiring a guarantee'.[20]

Thomson notes the similarity to lottery cases in epistemology.[21] Statistical evidence about lotteries can highly probabilify the claim that a specific ticket lost, but does not guarantee it. Lottery results printed in the newspaper, by contrast, can guarantee the ticket lost. As with individualised evidence of a litigated fact, newspaper evidence can, in Thomson's view, guarantee the ticket lost even if the likelihood of false belief from newspaper evidence is higher than the likelihood of false belief via merely statistical evidence.[22]

Thomson's suggestion has virtues. Statistical evidence, at least in the above kinds of cases, characteristically lacks a causal relation to the disputed claim. And statistical evidence

---

[16] For recent advances in statistical approaches to legal evidence, see Pardo (2013); Schneps and Colmez (2013); Koehler (2001); Kaye (2009); Roth (2010); Slatkin, Song and Murphy (2009); Tribe (1971); Goldberg (2013).

[17] Di Bello (2013) discusses the relationship between the claims that standards of proof are quantitative and that purely statistical evidence satisfies the standard.

[18] Thomson (1986).

[19] Thomson (1986: 206, 208–9, 214).

[20] Thomson (1986: 214).

[21] Thomson (1986: 207–208). Discussion of lottery cases in epistemology stems from Kyburg (1961). See also Nelkin (2000) and Hawthorne (2004).

[22] This can happen if, for example, the newspaper-reported lottery is smaller and there is some chance the newspaper misprinted the numbers.

plausibly fails to provide a guarantee: even excellent statistical evidence is straightforwardly consistent with the target claim's falsity. Individualised evidence, by contrast, typically has a causal relation and can characteristically – in some sense – generate a guarantee. If a person is convicted even though no evidence is causally related to the putative crime, furthermore, this seems inconsistent with justice; the defendant's crime has no bearing on whether he is convicted.[23]

But there are weaknesses in Thomson's account. Firstly, absent a more precise conception of causation, the account remains schematic. Causation is a notoriously 'murky', elastic notion, and Thomson does not specify the causal relation her account employs.[24]

Some base rates are plausibly causally related to some litigated claims, furthermore, which suggests that lack of causal relation fails to capture the inadequacy of merely statistical evidence. If local gang membership rates are sufficiently high, for example, this base rate can causally contribute to an individual participating in an illegal gang initiation rite. Since it is causal, the base rate is a candidate for forward-looking evidence, but nonetheless is statistical. The statistical evidence in the uncertain paternity case above is plausibly caused by paternity; if so, this is backward-looking causal evidence that is nonetheless statistical.[25] Base rates can be causally related to a common cause. Perhaps in some cases religious tenets causally contribute both to rates of domestic violence within that religion and a particular instance of domestic violence. Despite these causal relations, base rate evidence seems illegitimate for satisfying the burden of proof and is (arguably) unindividualised.

Realistic social base rate evidence does not typically generate high statistical likelihood (and tends to be complex). In the following fictional examples base rate statistics are high and so are a potential candidate for sufficient evidence.[26]

> First Gatecrasher. A music venue sells seats at its event, but does not issue, or record who purchases, tickets. Fern notices the gate is open, and decides to gatecrash. Other people see her gatecrashing and decide to follow. The managers realise only 10 seats were sold but 80 people attended, and summon the police. Fern happens to be one of the people apprehended. The statistical evidence against Fern is deemed sufficient to satisfy the burden of proof, since it is overwhelmingly likely she gatecrashed.

> Opportunistic Gatecrasher. A music venue sells seats at its event, but does not issue, or record who purchases tickets. One day the gate is left open and, taking advantage of the lack of ticketing system, many people gatecrash. Oppy walks past the venue and realises many people are gatecrashing. Oppy sees this as opportunity to gatecrash, and so joins in. The managers realise only 10 seats were sold, but

---

23 I sometimes, for concision, use terminology specific to one kind of legal burden, such as criminal trials, even though the ideas apply equally to other burdens of proof.

24 Enoch et al. (2012); Enoch and Fisher (2013); Redmayne (2008); and Blome-Tillmann (2015) raise similar concerns. The term 'murky' appears in Redmayne (2008: 300).

25 For more on DNA evidence and the distinction between statistical and individualised evidence, see Di Bello (2013); Pritchard (forthcoming); Stein (2005); Kaye and Sensabaugh (2000).

26 The first gatecrasher and opportunistic gatecrasher cases are from Blome-Tillmann (2015). Blome-Tillmann mistakenly claims Thomson's account is not vulnerable to the opportunistic gatecrasher case (Blome-Tillmann 2015: 110). This is because Blome-Tillmann overlooks forward-looking evidence in Thomson's account. Blome-Tillmann's account of Thomson's view cannot be correct, since it omits forward-looking individualised evidence such as evidence of motive.

80 people attended. Oppy is one of those arrested. The statistical evidence against Oppy is deemed sufficient to satisfy the burden of proof, since it is overwhelmingly likely he gatecrashed.

In these cases, arguably the statistical evidence is causally related to the person's guilt, yet is not individualised, does not generate a guarantee, and should not satisfy a burden of proof. But according to the *letter* of Thomson's account, the statistical evidence, by being causally related, qualifies as individualised. Since the statistics render guilt likely, Thomson's account cannot explain why they do not suffice for conviction. This objection is pressed in Blome-Tillmann (2015).

Responses are available. By refining the account of causation, one might deny Fern caused the statistic. There were, after all, many intervening causes. It is implausible, however, that an account of causation is available where gambling debt causes the theft but the high proportion of gatecrashers does not cause Oppy's opportunistic gatecrashing. A second response posits the way the statistic causes Oppy's crime, by providing motivation or cover, is importantly disconnected from how it epistemically supports guilt, namely by establishing a base rate likelihood. Perhaps, then, the case does not undermine the *spirit* of Thomson's proposal. Thirdly, advocates of a causal account of the epistemic difference between statistical and individualised evidence might argue the statistical evidence against Oppy is not itself causal, perhaps because it is abstract. For the evidence to be causal one must incorporate that Oppy notices the high proportion of gatecrashers, which motivates him. But then, the response continues, the evidence qualifies as causal precisely because one includes an individualised, not purely statistical, aspect.

An alternative line of response argues that the existence of cases illustrating that *sometimes* brute statistical evidence is causally related to the target fact does not undermine Thomson's account; Thomson identifies why brute statistical evidence is *characteristically* inadequate for conviction. Thomson's account explains that merely statistical evidence characteristically cannot guarantee the disputed fact, but individualised evidence – evidence with appropriate causal relations to the fact – characteristically can. Thus a refined version of Thomson's causal account may avoid these criticisms.

This response raises a second set of concerns about Thomson's account. It is unclear how causal relations generate a guarantee, or indeed what Thomson means by 'guarantee'. A witness might testify that although she thinks the perpetrator wore sandals, she cannot recall the incident well. A causal relationship obtains – the witness believes this because the perpetrator wore sandals – but the evidence cannot provide a guarantee. We might speculate that if Jones stole, his debt was a cause. But we are unsure whether he stole and, if so, whether debt was a motive. These pieces of evidence offer no guarantee. It is unclear from Thomson's account why combining evidence of this kind is epistemically or legally superior to statistical evidence. Perhaps lottery-style, statistical evidence provides a greater guarantee than combining many unreliable, but individualised, considerations.

To further cleave guarantees from causally connected evidence, note there can be non-causal facts that generate guarantees. These might include facts of legal definition, such as if the accused performed certain actions he thereby committed an offence, or medical diagnosis, such as if the witness has specific traits then he has a particular disorder. A lawyer might

6

articulate entailment relations such as if three people conspired as alleged, then at least three people knew the plan in advance. Evidence can include mathematical testimony from expert witnesses. Cold hit DNA evidence is typically characterised as statistical and, if the DNA is unmixed and non-degraded, produces a virtual certainty.[27] These examples of non-causal evidence do not, by themselves, satisfy a burden of proof. The prosecutor must determine, for example, that the DNA would not have been at the crime scene were the defendant innocent. But although DNA evidence has a causal element, the statistical reasoning guarantees some facts of the case, such as whether the accused's DNA was found at the scene. These examples illustrate that non-causal evidence can provide guarantees during a trial, which raises the question of whether and how causal relations generate a distinctive guarantee.

Perhaps the guarantee arises because causal relations can only obtain if the crime occurred. Jones's committing the crime did not cause eyewitness beliefs, for example, unless he committed the crime. But this is a trivial sense of guarantee and cannot be what Thomson intended. It is divorced from our epistemic state; fact finders do not know whether the causal relation between the evidence and the litigated fact holds unless they know whether the litigated fact obtains. So the logic of causation does not provide reasoners with a guarantee. The relation holds objectively, and is not one to which we have independent epistemic access.

To develop Thomson's account, more must be said about the relation between individualised evidence and guarantees. In section five I articulate one way we might understand this connection.

As noted above, a causal relation does not obtain unless the relata exist. This generates a further problem for Thomson's account.[28] If individualised evidence requires a causal relation, how is individualised evidence possible when the putative fact does not obtain or where the evidence misleads? If the defendant is innocent, his (putative) guilt cannot stand in any causal relationship to evidence. But then, according to Thomson's account, there is no individualised evidence. This is problematic for at least two reasons. Firstly, evidence can be individualised, as opposed to merely statistical, regardless of whether it misleads and whether the accused is innocent. Secondly, it is a desideratum of an account of legal evidence that – if evidence is compelling but misleading – the burden can be satisfied even if the judgement is false. A police officer can have legitimate reasonable suspicion, for example, when no crime is afoot and an innocent defendant can appear guilty beyond reasonable doubt. The causal account owes an explanation of how misleading evidence can be individualised.

---

[27] Devik (2006); Schneps and Colmez (2013: 69). If the test is performed correctly and if – as geneticists contend – the locations of the 13 peaks in genetic code are mutually independent, the probability that two people share a DNA match is 1/400 trillion. Enoch and Fisher (2013) also note some probative evidence is non-causal.

[28] Blome-Tillmann (2015) raises a version of this worry.

To conclude this section I articulate two further worries concerning Thomson's account of the connections amongst individualised, causally-related, and guarantee-providing evidence. Thomson holds evidence is individualised only when uniquely identifying. She writes,[29]

> Mrs. Smith believes she saw a one-legged, left-handed, entirely bald, and extremely tall man kill Bloggs. That is individualized evidence that a man with those four features killed Bloggs, for the (putative) fact that a man with those four features killed Bloggs would causally explain Mrs. Smith's believing she saw a man with those four features kill Bloggs… There are [99 other men in the world in addition to Mullins] who have all four features, so *getting individualized evidence against Mullins requires getting some fact in respect of which an appropriate causal role is played by a feature which distinguishes Mullins from [all other men in the world]*.
>
> In the revised version of the case, Mrs. Smith believes she saw a one-legged, left-handed, entirely bald, extremely tall, and one-eyed man kill Bloggs… [O]ur further evidence suggests that only Mullins has all five features… To the extent to which we believe that further evidence, then, we shall take ourselves to have individualized evidence, not merely that a man with those five features killed Bloggs, but that Mullins did – he being the only available candidate with the five features.

The italicised condition would help underwrite a guarantee of guilt. But as a condition on individualised evidence, it is too strong. Evidence can be individualised and against the defendant without being uniquely identifying or conclusive. This point echoes Donnellan's and Kripke's objection to descriptivism as a semantics for proper names: we can have a *de re*, individualising thought about something without being able to uniquely discriminate it.[30] This underscores that individualised evidence is far from a guarantee of guilt, at least in the strong, ordinary sense of 'guarantee'.

Secondly, the notion of 'guarantee' arguably sits uncomfortably with lower standards. Thomson extends her account from criminal to civil burdens. According to Thomson, various strengths of standard of proof correspond to how sure fact finders are of having a guarantee of culpability. She writes,[31]

> Our law requires the jury in a criminal case to be sure beyond a reasonable doubt that the defendant is guilty before imposing liability on him; the friend of individualized evidence may be taken to say that the jury must be sure beyond a reasonable doubt that the defendant is guilty because of being sure beyond a reasonable doubt that there are facts available to it which guarantee that the defendant is guilty. Our law requires the jury in a case in tort to believe no more than that it is more probable than not that the defendant is guilty; the friend of individualized evidence may be taken to say that the jury must believe it is more probable than not that the defendant is guilty because of believing it more probable than not that there are facts available to it which guarantee that the defendant is guilty.

We might further extend this treatment: to satisfy a 'clear and compelling evidence' standard fact finders must believe they have clear and compelling evidence that that there are facts available to them which guarantee the defendant is liable; to satisfy a 'reasonable suspicion' standard the officer must believe she has reasonable suspicion that that there are facts available to her which guarantee that crime is afoot.

---

29   Thomson (1986: 217, emphasis mine).
30   Donnellan (1972); Kripke (1972). Thanks to Jon Garthoff for helpful discussion of these issues.
31   Thomson (1986: 215, 219).

But our ordinary notion of 'guarantee' seems inappropriate for understanding standards lower than beyond reasonable doubt. Plausibly finding fault in civil cases does not require any evidence or chance of a *guarantee* of liability, for example, instead simply requiring a certain level of epistemic support. Perhaps a court can appropriately find in favour of the plaintiff, in other words, even if no evidence purports to guarantee liability, but the preponderance of available evidence indicates liability.

Thomson's account of the inadequacy of statistical evidence appears unsuccessful as it stands. To render the view more compelling, we need a better understanding of the relevant notion of guarantee, the causal relation, and how causal relations underwrite guarantees.

4. Sensitivity and Incentives

An alternative view holds that merely statistical evidence cannot satisfy a legal burden because legal standards of proof require one's judgement to be sensitive to the truth of the litigated facts.[32] Enoch et al. (2012) and Enoch and Fisher (2013) argue the deficiency of merely statistical evidence is that it obtains regardless of the specific crime, and so a belief based on statistical evidence would be held even were the accused innocent. Individualised evidence, by contrast, characteristically engenders sensitive beliefs; if the defendant did not commit the crime, the individualised evidence would not obtain.[33]

> Sensitivity of Beliefs. S's belief that p is sensitive iff had it not been the case that p, S (likely) would not have believed p.

Extending the notion of sensitivity of beliefs suggests a corresponding 'sensitivity of evidence': Evidence is sensitive to a claim iff (roughly speaking) were the claim false, the evidence would likely not obtain.[34]

Enoch et al.'s suggestion enjoys a great deal of plausibility. It is plausibly a principle of justice and good reasoning that a person should not be found culpable unless evidence used to convict is counterfactually dependent on the crime. The conviction should be sensitive to the transgression. The court should be able to assert that if the accused did not the commit the crime, he would not have been convicted. Thomson's suggestion employs the causation relation to unpack this idea; the sensitivity approach employs modal conditions.

Enoch et al. argue, however, that sensitivity as an epistemic value does not have legal value and courts should not care about the sensitivity of their judgements.[35] Instead they advance a non-epistemic vindication of the legal value of sensitivity, and thus a practical vindication of the distinction between individualised and statistical evidence. They argue that relying on statistical evidence undermines incentives to obey the law. If evidence sufficient for finding culpable obtains regardless of whether the person transgresses, Enoch et al. reason, the

---

[32] For information on sensitivity see Nozick (1981) and DeRose (1996). For a sensitivity-based approach to lottery cases in epistemology see Dretske (1971) and DeRose (forthcoming: chapter 5).

[33] At least, those beliefs are sensitive in the good case; see discussion below.

[34] Here the evidence must be understood with some imprecision. It is not that 70 people gatecrashed, for example, it is that many people gatecrashed. See also Blome-Tillmann (2015).

[35] Enoch et al. (2013: 211–215); Enoch and Fisher (2013: 577–581).

person might as well transgress because 'whatever he decides will have negligible influence on the likelihood of his being punished'.[36] If purely statistical evidence suffices to satisfy the burden of proof, then a person deciding whether to purchase a ticket or gatecrash, for example, has no incentive (stemming from evidence law) to purchase the ticket. Similarly, taxi drivers lose their incentive to drive cautiously.

I think this incentive-based explanation of the need for individualised evidence is mistaken. Firstly, consider the gatecrasher case, which is the example they select to illustrate their claims about incentive structures. In this case there remains an incentive to obey the law by not entering the venue; by not entering, the person avoids arrest. A more significant objection is that cases with the appropriate incentive structure are rare. Enoch at al.'s explanation only applies when the crime occurs regardless of whether the person participates, the person is already in the class of suspects to be investigated and will remain so regardless of whether he participates, the person must decide whether to participate in the law-breaking, *and* an investigation or other legal action will occur regardless of whether the person participates. Cases with this structure might include activities such as small-fry tax fraud, littering, and television licence evasion. If defendants for these infractions were identified *and* prosecuted with purely statistical evidence, the incentive from legal evidence would arguably be diminished.[37] But typically decisions about whether to transgress concern cases where the crime or harm to be investigated would not otherwise occur, and so no investigation would occur but for the person's transgressing. This includes almost all thefts, assaults, harassment, negligent accidents, embezzlement etc. So even with merely statistical evidence, incentive is provided: if the person had not transgressed, there would be no investigation or risk of punishment.

Enoch et al. acknowledge this concern. They write,[38]

> The incentive story thus has different implications across the two categories of cases: those in which the act would likely be performed by others regardless of whether the would-be perpetrator decides not to engage in it, and personal-context cases [such as spousal abuse], in which the act will not likely be carried out by anyone else. In the latter type of case, the statistical evidence against the defendant ought to be admissible at trial.

What Enoch et al. fail to appreciate, I think, is that 'the latter type of case' includes most crimes and liabilities. If Jones does not coldcock Williams on his way home, probably no one else will. Their incentive-based account, rather than vindicating the legal distinction between individualised and merely statistical evidence, instead vindicates the adequacy of statistical evidence in almost every instance.

---

[36] Enoch and Fisher (2013: 583).

[37] Although I do not have space to develop this claim here, I believe Enoch et al. elide the distinction between statistical evidence being used to identify and convict the accused. If statistical evidence were used to identify suspects from the general population to prosecute, this *might* reduce incentives (from some aspects of evidence law) to not transgress. (Although even here, for reasons articulated below, I think the incentive-based account cannot capture the nature of the wrong.) But the use of statistics proposed is *convicting* the suspect once identified. Given this, the incentive-based account is less successful.

[38] Enoch and Fisher (2013: 608); see also Enoch et al. (2012: 219).

Allowing purely statistical evidence to satisfy the burden, furthermore, might well disincentivise crime: It lifts a restriction on admissible evidence, and thereby renders conviction more likely.

Arguably the strongest case for Enoch et al.'s incentive-based account concerns drivers employed by the smaller taxi service. Drivers of green taxis will not be found liable in any case adjudicated solely by market share evidence. So perhaps they lose incentive (from legal evidence) if purely statistical evidence can satisfy a burden. But even in this instance the incentive-based account is uncompelling. To see why, note the difference between statistical evidence sufficing for a finding and statistical evidence being the only admissible evidence. If it were the only admissible evidence, green taxi drivers would have no incentive (from legal evidence) to drive cautiously: they would never be found liable. But since – regardless of whether statistical evidence suffices – other kinds of evidence are admissible, drivers retain incentives from legal evidence: for any particular accident, there is likely to be particularised incriminating evidence. Only in unusual cases would the court rely wholly on market share evidence.

A further weakness of the incentive-based account is that even if merely statistical evidence is employed persons retain incentives from other sources, such as the motivation to not steal, feel guilty, be injured, or injure others. Enoch et al. do not mention these other incentives, but they dilute the importance of incentives stemming from legal evidence.

Finally, Enoch et al.'s proposal, which concerns influencing one aspect of citizens' incentive structures in some kinds of cases, does not provide adequate warrant for our response to the prospect of satisfying the burden of proof with purely statistical evidence. When we consider the cases described in section two, there seems something *unjust* about convicting, finding liable, arresting, searching, or detaining on purely statistical evidence. Plausibly those people are *wronged*. But Enoch at al.'s incentive-based account, which is rooted in practical policy considerations, is not the right kind of explanation to vindicate those moral reactions.

We can set aside Enoch et al.'s incentive-based explication of the distinction between individualised and statistical evidence. The question remains whether a sensitivity-based account explains the inadequacy of statistical evidence. As noted above, it seems remiss for a court to convict unless they would have acquitted were the defendant innocent. Plausibly this sensitivity-based epistemic condition itself has legal value. The idea merits investigation.

The sensitivity account raises some similar concerns as the causal account. Some statistical evidence – or beliefs based on statistical evidence – is sensitive to the crime's occurring, for example. Recall Fern the first gatecrasher and Oppy the opportunistic gatecrasher. In these cases, statistical evidence generates the relevant sensitivity counterfactuals. If Fern did not gatecrash, the statistical evidence would likely not have obtained. If Oppy had not gatecrashed, the courts would not have found him liable.[39]

---

[39] Blome-Tillmann (2015: 106–107). Blome-Tillman holds that the latter counterfactual is true since had Oppy not gatecrashed he would not have entered the venue. Many thanks to Martin Smith for helpful comments on these issues.

Some of the same responses are available: Like advocates of the causal account, advocates of the sensitivity account might argue the account identifies what is characteristically inadequate about employing statistical evidence to satisfy a burden, and the account is thereby consistent with some marginal instances in which statistical evidence generates sensitive beliefs. In response note that, assuming Tom is the father, statistical evidence in the uncertainty paternity case generates sensitive beliefs, and underwrites a sufficiently high likelihood, yet is inadequate for finding liability. This suggests that mundane, central kinds of statistical evidence engender sensitive beliefs.[40]

A further problem for the proposal is that much individualised evidence is insensitive: regardless of whether the accused is culpable, there is incriminating individualised evidence against him. This can occur with misleading or inconclusive evidence. Perhaps, for example, the accused's fingerprints are at the scene because he is an acquaintance of the victim. The fingerprints can feature in the prosecution's narrative as individualised evidence, but may not be sensitive to the crime.

Note too that sensitivity is factive; a belief cannot be sensitive unless true. This creates a challenge for a sensitivity account of the legal value of individualised evidence. To see why, consider a case in which the accused is innocent but there is compelling, misleading, incriminating, individualised evidence. To determine whether the fact finder's belief is sensitive we must ask, 'Were the accused innocent, would the fact finder believe he is innocent?' Since the defendant is actually innocent, the answer is no. No false conviction is sensitive. If the burden of proof requires that the judgement is sensitive, as the sensitivity account suggests, no false conviction satisfies the burden. But, as noted above, it is a desideratum of legal epistemology that some false convictions satisfy the burden.[41]

Enoch et al. respond to these kinds of objections by arguing sensitivity is a hallmark of *good* individualised evidence, not just any individualised evidence; statistical evidence, by contrast, is insensitive even when it is good statistical evidence.[42] But it bears noting that as an account of the characteristic virtue of individualised evidence over statistical evidence, it is a weakness of the sensitivity account that much individualised evidence is insensitive. In section five I explore whether a relation known as 'normic support', or a kindred relation, is a better hallmark of individualised evidence and can better explain its legal and epistemic value.

5. Normic Support
Many epistemologists hold that epistemic support concerns probabilification.[43] They hold that a body of evidence E epistemically supports proposition p iff p is likely, given E, to be true; a body of evidence E_1 epistemically supports proposition p more than E_2 does iff p is more likely given E_1 than given E_2.

Martin Smith (2010) challenges this orthodoxy. He notes that in the lottery case, purely statistical evidence can establish that it is extremely likely a particular ticket lost, but cannot justify outright belief that the ticket lost. Reading lottery results in a newspaper, on the other hand, can justify outright belief. Individualised evidence can justify belief even if the claim is more likely to be false than if it were based on purely statistical evidence.

Smith argues that individualised evidence, but not statistical evidence, 'normically supports' the target claim. Evidence normically supports a conclusion when, roughly speaking, given that evidence, p would normally be true. This notion of normalcy does not reduce to statistical frequency.[44] If the evidence obtains yet p is false, some abnormality or malfunction has occurred. The error demands explanation. When evidence statistically indicates p but p is false, by contrast, it is not really an error; it is simply a case of 'you win some you lose some'.[45]

Smith employs a possible worlds framework for understanding normic support. He writes,[46]

> A body of evidence E normically supports a proposition p just in case p is true in all the most normal worlds in which E is true. Alternately, E normically supports p just in case the most normal worlds in which E is true and p is true are more normal than any world in which E is true and p is false… A body of evidence E normically supports a proposition p *more strongly* than it normically supports a proposition q just in case the most normal worlds in which E is true and q is false are more normal than any world in which E is true and p is false.

There are concerns with this account of normic support and, in particular, the possible world framework proposed. One concern is that more normal claims receive more normically-supportive evidence *just in virtue of* being normal. If p is adequately normal an intuitively irrelevant fact F will qualify as normically supporting p. Suppose p obtains in almost all normal words (p is 'grass is green', for example). Now consider an intuitively irrelevant fact F, such as 'polar bear hind paws are elongated'. The normal F-worlds are p-worlds; the worlds where F is true and p is true are more normal than any world where F is true and p is false. This is because not-p worlds are almost all abnormal. And so the seemingly irrelevant fact F normically supports p. A related worry – one common to possible world analyses – concerns logical truths. Logical truth T, which obtains in all worlds and so obtains in all normal worlds, is normically supported by an irrelevant fact F, because normal F-worlds are T-worlds.

Even if Smith's understanding of the normic support relation is not correct, perhaps an epistemic support relation similar to Smith's normic support can vindicate the distinction between individualised and merely statistical evidence.[47] In what follows I use 'normic support' to denote epistemic support relations that share key features with Smith's proposal but may vary in the details. (In particular one might reject analysing normic support in terms

---

[44]   Smith (2010: 16).

[45]   The expression 'you win some you lose some' appears in Enoch et al. (2012: 208).

[46]   Smith (ms: 20–21). See also Smith (2010:16–17). The actual world may not be maximally normal; misleading individualised evidence is possible.

[47]   Smith (ms: 41ff.) applies normic support to legal epistemology.

of possible worlds.[48]) Consider the first five vignettes in section two. Reflecting on what normally follows from the evidence plausibly illuminates various judgements about the cases. Insofar as it might seem permissible to search partying underage undergraduates, for example, this is because given the evidence they would *normally* have alcohol and their not having alcohol would be an abnormal, surprising fact. This might obtain, for instance, if underage partygoers exhibit hallmarks of drinking, such as vomiting in bushes. If a police search is conducted purely on statistical grounds, by contrast, it seems illegitimate. The uncertain paternity vignette stipulates the samples were mixed. Given this laboratory error, it is normal and does not demand further explanation if the results are misleading. And so we cannot rely on those results in court. If the tests were conducted with non-mixed DNA, by contrast, the match would normally support the paternity claim since error would demand explanation. Perhaps a normic support relation, then, can vindicate the epistemic and legal distinction between individualised and statistical evidence.

If Smith's normic support, or a kindred relation, is a condition on satisfying burdens of proof, this may illuminate some ethical concerns about 'stop and frisk' tactics. Suppose many members of a demographic commit a particular crime. For any individual, given only demographic evidence, it is statistically relatively likely they commit the crime. And now suppose the police frisk an individual based solely on demographic evidence. If burdens of proof require only statistical likelihoods, the officer's frisk would convey, 'I judge (with reasonable suspicion, given only demographic evidence) it is statistically relatively likely the person committed the crime.' Given the stipulation, the judgement could be true and well-grounded; arguably the judgement does not capture the wrong of the frisk. If instead the burden of proof concerns normalcy, the officer's frisk conveys, 'I judge (with reasonable suspicion, given the evidence) it is normal that the person committed the crime; if the person has not committed the crime it is a departure from normalcy'. Plausibly this judgement, unlike the first, is a condemnation of the individual or their group. It is wrong to judge with only demographic evidence that someone would normally transgress. Thus part of the moral seriousness of 'stop and frisk' may stem from the nature of the burden of proof.

Appealing to the idea of normalcy may illuminate promising aspects of the causal and sensitivity accounts. Causal relations often generate normic support relations. If an accident causes eyewitness evidence indicating the taxi was red, for example, normally the taxi is red. If the liable taxi is not red, this is abnormal and demands explanation. Evidence that normally supports a conclusion typically underwrites the belief's sensitivity. Enoch et al. note that sensitivity and normic support often correlate, but remain agnostic about which is explanatorily more basic. They suggest normic support is less basic because the notion of

---

[48] For alternative discussions of normalcy, see Pettit (1999); Millikan (1984); Nickel (2008); Schurz (2001); Haslanger (2014); Morreau (1997); Spohn (2014); and Hauska (2008). Gardiner (2015) discusses the relationship between possible worlds and normalcy. Normalcy is not usually discussed by philosophers as a subject matter, but instead the notion of normalcy is used to illuminate other areas, such as character in ethical theory, populations in philosophy of biology, illness in philosophy of disability, conditionals in metaphysics, ceteris paribus laws in philosophy of science, generics in philosophy of language, and in the semantics for nonmonotonic logic. I am very grateful to Bernhard Nickel, Liz Camp, and Martin Smith for helpful insights on these topics.

'which errors call for explanation' seems unilluminating and opaque.[49] Enoch et al. are correct that 'what calls for explanation' is a poor candidate for an explanatory foundation. But this is not the heart of normic support; it is simply a characteristic feature. Normic support is a relation of epistemic support; plausibly it generates facts about which beliefs are sensitive and about which errors demand explanation. If a belief is insensitive, in other words, plausibly this is because the evidence fails to normically support the belief. Normic support seems the more basic notion.[50]

Thomson argues that an epistemically valuable feature of individualised evidence is that it – unlike statistical evidence – can provide a guarantee. I noted above this notion was underdescribed in Thomson's account. Perhaps a normic support relation can illuminate this 'guarantee'. If evidence normically supports p, and yet not p, something is amiss and is not as it seems. The error demands explanation because something abnormal has occurred. Perhaps this is the kind of guarantee demanded by the burden of proof.

Enoch et al. demur. They write,[51]

> Why should the law especially care about avoiding mistakes that call for explanation? Mistakes that do not call for explanation seem – absent some storytelling otherwise, at least – just as harmful to the relevant party, just as detrimental to the relevant social interests, and so on, as mistakes that do call for explanation.

But this point seems mistaken. If a court convicts a defendant erroneously, and the only response available is 'you win some you lose some' because the court relied on statistical evidence to satisfy the burden, the court has wronged the defendant. If a person is convicted, found liable, searched, or arrested without participating in the alleged activity, this error ought arise from some abnormal feature. Plausibly justice – also public trust in the legal system – demands this.

Smith's account of the relationship between error and normic support, if correct, can further illuminate Thomson's guarantee. He writes,[52]

> If one believes that a proposition p is true, based upon evidence that normally supports it then, while one's belief is not assured to be true, this much is assured: If one's belief turns out to be false, then the error has to be *explicable* in terms of disobliging environmental conditions, deceit, cognitive or perceptual malfunction or some other interfering factor. In short, the error must be attributable to *mitigating circumstances* and thus *excusable*, after a fashion. Errors that do not fall into this category are naturally regarded as errors for which one must bear *full responsibility* – errors for which there is no excuse.

According to Smith, if one's evidence normically supports p, and one believes p based on that evidence, the person is not responsible for the error if p is false. I am doubtful of Smith's claim concerning epistemic responsibility, since it seems evidence could normically

---

[49] Enoch et al. (2012: 210).
[50] Gardiner (2017b) develops the claim that epistemic modal conditions supervene on other epistemic facts, and are explanatorily less basic than those facts.
[51] Enoch et al. (2013: 214).
[52] Smith (ms: 18, emphasis in original).

support p, p not obtain, yet the error accrue to the person's reasoning and be something the reasoner is responsible for. Suppose, for example, Oliver believes p, where p is 'Jones murdered Jill'. He believes this because Jones's clothes are bloodstained. Oliver's belief is thus based upon evidence that normically supports p. But Jones believes p because he is confident that only butchers wear bloody clothes and butchers are angry and murderous. In this case, the evidence normically supports p, and Oliver believes p because of that evidence, but the error nonetheless accrues to Oliver's poor reasoning. If Oliver had reasoned better, he may well have realised Jones's innocence.

Even if Smith's claim about the relationship between normic support and epistemic responsibility is incorrect, a normalcy-based epistemic support relation might plausibly illuminate the sense of 'guarantee' generated by individualised evidence: When a conclusion normally follows from the evidence, perhaps the evidence does not merely probabilify the conclusion, it also entails that if the evidence is true, and yet the conclusion false, something is amiss and abnormal. Perhaps this approaches the kind of guarantee demanded by legal burdens of proof.[53]

Bibliography
Armour, Jody (1994) 'Race Ipsa Loquitur: Of Reasonable Racists, Intelligent Bayesians, and Involuntary Negrophobes' *Stanford Law Review* 46(4): 781–816.
Bacigal, Ronald (2004) 'Making the Right Gamble: The Odds on Probable Cause' *Mississippi Law Journal* 74: 279–340.
Blome-Tillmann, Michael (2015) 'Sensitivity, Causality, and Statistical Evidence in Courts of Law' *Thought* 4(2): 102–112.
Buchak, Lara (2014) 'Belief, Credence, and Norms' *Philosophical Studies* 169(2): 285–311.
Cohen, Jonathan (1977) *The Probable and the Provable* Oxford University Press.
Colyvan, Mark, Helen Regan and Scott Ferson (2001) 'Is it a Crime to Belong to a Reference Class?' *Journal of Political Philosophy* 9(2): 168–181.
DeRose, Keith (1996) 'Knowledge, Assertion, and Lotteries' *Australasian Journal of Philosophy* 74: 568–580.
DeRose, Keith (forthcoming) *The Appearance of Ignorance: Knowledge, Skepticism, and Context, Vol. 2* Oxford University Press.
Devik, Keith (2006) 'Statisticians Not Wanted' blog post https://www.maa.org/external_archive/devlin/devlin_09_06.html (accessed March 2016).
Di Bello, Marcello (2013*) Statistics And Probability in Criminal Trials: The Good, the Bad and the Ugly* Doctoral Dissertation, Stanford University.

Donnellan, Keith (1972) 'Proper Names and Identifying Descriptions' in Donald Davidson and Gilbert Harman (eds.) *Semantics of Natural Language* Dordrecht: D. Reidel, pp. 356–379.

Dretske, Fred (1971) 'Conclusive Reasons' *Australasian Journal of Philosophy* 49: 1–22.

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

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

Franklin, James (2006) 'Case Comment – United States v. Copeland, 369 F. Supp. 2d 275 (E.D.N.Y. 2005): Quantification of the 'Proof Beyond Reasonable Doubt' Standard' *Law, Probability, and Risk* 5: 159–165.

Gardiner, Georgi (2015) 'Normalcy and the Contents of Philosophical Judgements' *Inquiry* 58(7): 700–740.

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

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

Gardiner, Georgi (forthcoming) 'Evidentialism and Moral Encroachment' *Believing in Accordance with the Evidence: New Essays on Evidentialism* Kevin McCain (ed.) Springer.

Goldberg, Erica (2013) 'Getting Beyond Intuition in the Probable Cause Inquiry' *Lewis & Clark Law Review* 17: 789–850.

Haslanger, Sally (2014) 'The Normal, the Natural and the Good: Generics and Ideology' *Politica & Società* 3: 365–392.

Hauska, Jan (2008) 'Dispositions and Normal Conditions' *Philosophical Studies* 139: 219–232.

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

Kaye, David (2009) 'Rounding up the Usual Suspects: A Legal and Logical Analysis of DNA Trawling Cases' *North Carolina Law Review* 87: 425–503.

Kaye, David and George Sensabaugh (2011) 'Reference Guide on DNA Identification Evidence' in *Reference Manual on Scientific Evidence, 3rd Edition* Federal Judicial Center, pp 129–210.

Kerr, Orin (2012) 'Why Courts Should Not Quantify Probable Cause' in Michael Klarman, David Skeel, Carol Steiker (eds.) *The Political Heart of Criminal Procedure: Essays on Themes of William J. Stuntz* Cambridge University Press, pp. 131–143.

Koehler, Jonathan (2001) 'When Do Courts Think Base Rate Statistics are Relevant?' *Jurimetrics Journal* 42: 373–402.

Kripke, Saul (1972) *Naming and Necessity* Harvard University Press.

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

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

Laudan, Larry (2012) 'Is it Finally Time to Put "Proof Beyond a Reasonable Doubt" Out to Pasture?' in Andrei Marmor (ed.) *Routledge Companion To Philosophy Of Law* New York: Routledge, pp. 317–332

Lerner, Craig (2003) 'The Reasonableness of Probable Cause' *Tex. Law Review* 81: 951–1029.

Lippke, Richard (2010) 'Punishing the Guilty, Not Punishing the Innocent' *Journal of Moral Philosophy* 7(4): 462–88.

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

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

Millikan, Ruth (1984) *Language, Thought, and Other Biological Categories: New Foundations for Realism* MIT Press.

Morreau, Michael (1997) 'Fainthearted Conditionals' *Journal of Philosophy* 94(4): 187–211.

Mulrine, Thomas (1997) 'Reasonable Doubt: How in the World Is It Defined?' *American University Journal of International Law and Policy* 12(1): 195–225.

Nelkin, Dana (2000). 'The Lottery Paradox, Knowledge, and Rationality' *Philosophical Review* 109(3): 373–409.

Nesson, Charles (1979) 'Reasonable Doubt and Permissive Inferences: The Value of Complexity' *Harvard Law Review* 92(6): 1187–1225.

Nesson, Charles (1985) 'The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts' *Harvard Law Review* 98: 1357–1392.

Newman, Jon (2006) 'Quantifying the Standard of Proof Beyond a Reasonable Doubt: A Comment on Three Comments' *Law, Probability, and Risk* 5: 267–269

Nickel, Bernhard (2008) 'Generics and the Ways of Normality' *Linguistics and Philosophy* 31(6): 629–648.

Nozick, Robert (1981) *Philosophical Explanations* Oxford University Press.

Pardo, M.S. (2013) 'The Nature and Purpose of Evidence Theory' *Vanderbilt Law Review* 66: 547–613.

Pardo, M.S. and R.J. Allen (2008) 'Juridical Proof and the Best Explanation' *Law and Philosophy* 27: 223–268.

Pettit, Philip (1999) 'A Theory of Normal and Ideal Conditions' *Philosophical Studies* 96: 21–44.

Pritchard, Duncan (forthcoming) 'Legal Risk, Legal Evidence and the Arithmetic of Criminal Justice' *Jurisprudence*.

Redmayne, Mike (2008) 'Exploring the Proof Paradoxes' *Legal Theory* 14: 281–309.

Risinger, D. Michael (2010) 'Tragic Consequences of Deadly Dilemmas: A Response to Allen and Laudan' *Seton Hall Law Review* 40(3): 991–1021

Roth, Andrea (2010) 'Safety in Numbers? Deciding When DNA Alone is Enough to Convict' *New York University Law Review* 85: 1130–1185.

Sand, Leonard and Danielle Rose (2003) 'Proof Beyond All Possible Doubt: Is There a Need for a Higher Burden of Proof when the Sentence may be Death?' *Chicago-Kent Law Review* 78: 1359–1375.

Schmalbeck, Richard (1986) 'The Trouble with Statistical Evidence' *Law and Contemporary Problems* 49(3): 221–236.

Schneps, Leila and Coralie Colmez (2013) *Math on Trial* Basic Books.

Schurz, Gerhard (2001) 'What is "Normal"? An Evolution-Theoretic Foundation for Normic Laws and Their Relation to Statistical Normality' *Philosophy of Science* 68(4): 476–497.

Simon, Rita James (1969) 'Judges' Translations of Burdens of Proof into Statements of Probability' in J. Kennelly and J. Chapman (eds.) *The Trial Lawyer's Guide* Mundelein: Callaghan, pp. 103–114.

Simon, Rita James and Linda Mahan (1971) 'Quantifying Burdens of Proof: A View from the Bench, the Jury, and the Classroom' *Law and Society Review* 5(3): 319–330.

Slatkin, Montgomery, Yun S. Song & Erin Murphy (2009) 'Average Probability that a "Cold Hit" in a DNA Database Search Results in an Erroneous Attribution' *Journal of Forensic Sciences* 54(1): 22–27.

Smith, Martin (2010) 'What Else Justification Could Be?' *Noûs* 44(1): 10–31.

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

Smith, Martin (ms) 'Justification, Normalcy and Evidential Probability' http://philpapers.org/rec/SMIJNA-2 (accessed March 2016).

Spohn, Wolfgang (2014) 'The Epistemic Account of Ceteris Paribus Conditions' *European Journal for the Philosophy of Science* 4(3): 385–408.

Staffel, Julia (forthcoming) 'Three Puzzles about Lotteries' in Igor Douven (ed.) *Lotteries, Knowledge, and Rational Belief: Essays on the Lottery Paradox* Cambridge University Press.

Stein, Alex, (2005) *Foundations of Evidence Law* Oxford University Press.

Thomson, Judith Jarvis (1986) 'Liability and Individualized Evidence' *Law and Contemporary Problems* 49(3): 199–219.

Tillers, Peter and Jonathan Gottfried (2006) 'Case comment – United States v. Copeland, 369 F. Supp. 2d275 (E.D.N.Y. 2005): A Collateral Attack on the Legal Maxim That Proof Beyond A Reasonable Doubt Is Unquantifiable?' *Law, Probability, and Risk* 5: 135–157.

Tribe, Laurence (1971) 'Trial By Mathematics: Precision and Ritual in the Legal Process' *Harvard Law Review* 84(6): 1329–1393.

Tversky, Amos and Daniel Kahneman (1982) 'Evidential Impact of Base Rates' in Daniel Kahneman, Paul Slovic and Amos Tversky (eds.) *Judgment Under Uncertainty: Heuristics and Biases* Cambridge University Press, pp. 153–162.

Vars, Frederick (2010) 'Toward a General Theory of Standards of Proof' *Catholic University Law Review* 60(1): 1–46.

Volokh, Alexander (1997) 'n Guilty Men' *University of Pennsylvania Law Review* 146: 173–216

Walen, Alec (2015) 'Proof Beyond a Reasonable Doubt: A Balanced Retributive Account' *Louisiana Law Review* 76(2): 355–446.

Weinstein, Jack and Ian Dewsbury (2006) 'Comment on the Meaning of "Proof Beyond a Reasonable Doubt"' *Law, Probability and Risk* 5: 167–173.