Profiling and Proof: Are Statistics Safe?*

GEORGI GARDINER

Abstract

Many theorists hold that outright verdicts based on bare statistical evidence are unwarranted. Bare statistical evidence may support high credence, on these views, but does not support outright belief or legal verdicts of culpability. The vignettes that constitute the lottery paradox and the proof paradox are marshalled to support this claim. Some theorists argue, furthermore, that examples of profiling also indicate that bare statistical evidence is insufficient for warranting outright verdicts.

I examine Pritchard's and Buchak's treatments of these three kinds of case. Pritchard argues that his safety condition explains the insufficiency of bare statistical evidence for outright verdicts in each of the three cases, while Buchak argues that her treatment of the distinction between credence and belief explains this. In these discussions the three kinds of cases – lottery, proof paradox, and profiling – are treated alike. The cases are taken to exhibit the same epistemic features. I identity significant overlooked epistemic differences amongst these three cases; these differences cast doubt on Pritchard's explanation of the insufficiency of bare statistical evidence for outright verdicts. Finally, I raise the question of whether we should aim for a unified explanation of the three paradoxes.

1. Introduction

This paper contrasts the 'proof paradox' in legal epistemology, Gilbert Harman's 'lottery paradox' and Lara Buchak's 'profiling paradox'.[1] All three 'paradoxes' comprise two vignettes which jointly suggest that bare statistical evidence is insufficient for outright judgement. The three pairs of vignettes – the lottery, proof, and profiling vignettes – are often treated as exhibiting identical epistemic features. But, I argue, the three pairs of vignettes exhibit important overlooked epistemological differences. Existent proof paradox vignettes, furthermore, do not themselves constitute a unified epistemological kind, and instead are different species of example that theorists unwittingly conflate.

I describe and cast doubt on Duncan Pritchard's response to the proof paradox. I employ 'stipulative epidemiology' to illustrate how

* Editors' note: this is the winning entry of the 2019 *Philosophy* Essay Prize.

[1] This is not Buchak's term.

doi:10.1017/S0031819120000030 © The Royal Institute of Philosophy, 2020

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

different causal structures can underwrite different modal properties; these differences can be obscured by shared patterns of statistical prevalence. Finally, I draw on recent theorising about moral encroachment to interrogate whether we should seek a unified explanation of the three paradoxes.

2. Safety and the Proof Paradox

The proof paradox in epistemology is constructed by contrasting two vignettes.[2]

> **Taxi One.** A vehicle hit Jeanette one evening. She could determine it was a taxi, but could not discern the colour. The Red Taxi company operates 90% of taxis in town. The remaining 10% are operated by the Green Taxi company. Using only the evidence described here, Jeanette sues the Red Taxi company.

[2] See Laurence Tribe 'Trial by Mathematics: Precision and Ritual in the Legal Process' *Harvard Law Review* 84 (1971) 1329–1393; Jonathan Cohen *The Probable and the Provable* (Oxford: Oxford University Press, 1977); Judith Jarvis Thomson 'Liability and Individualized Evidence' *Law and Contemporary Problems* 49 (1986) 199–219; Hock Lai Ho 'The Legal Concept of Evidence' *Stanford Encyclopedia of Philosophy* http://plato.stanford.edu/archives/win2015/entries/evidence-legal/ (Winter 2015); Mike Redmayne 'Exploring the Proof Paradoxes' *Legal Theory* 14 (2008) 281–309; Marcello Di Bello *Statistics and Probability in Criminal Trials: The Good, the Bad and the Ugly* Doctoral Dissertation, (Stanford University, 2013); David Enoch, Levi Spectre, and Talia Fisher 'Statistical Evidence, Sensitivity, and the Legal Value of Knowledge' *Philosophy and Public Affairs* 40 (2012) 197–224; Martin Smith 'When Does Evidence Suffice for Conviction?' *Mind* 127 (2018) 1193-1218; Martin Smith 'Accuracy, Proof and the Right Against Wrongful Conviction' (ms); Clayton Littlejohn 'Truth, Knowledge, and the Standard of Proof in Criminal Law' *Synthese* (2017); Renee Bolinger 'The Rational Impermissibility of Accepting (Some) Racial Generalizations' *Synthese* (2018); Georgi Gardiner 'Legal Burdens of Proof and Statistical Evidence' in David Coady and James Chase (eds.) *The Routledge Handbook of Applied Epistemology* (Routledge, 2018) 171–195; Michael Pardo 'The Paradoxes of Legal Proof: A Critical Guide' *Boston University Law Review* (2019); Julia Staffel 'Three Puzzles about Lotteries' in Igor Douven (ed.) *Lotteries, Knowledge, and Rational Belief: Essays on the Lottery Paradox* (Cambridge University Press, forthcoming). For a survey of proof paradox research, see Georgi Gardiner 'Legal Epistemology' in *Oxford Bibliographies: Philosophy*, ed. Duncan Pritchard (Oxford: Oxford University Press, 2019).

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

The judge reasons that it is more likely than not that the Red Taxi company is liable, and rules against them.

Taxi Two. A vehicle hit Jeanette one evening. 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 75% of the time. With this evidence Jeanette successfully sues the Red Taxi company.[3]

Plausibly in the former case the 'bare statistical evidence' licenses a relatively high credence – perhaps 0.90 – in the target claim. But there is something amiss with establishing legal liability on this evidence alone. In the second case, non-statistical, 'individualised' evidence is available. The overall probability of the disputed fact given the evidence might well be lower, yet the relevant burden of proof – preponderance of evidence – could well be satisfied on this basis. The central problem raised by the proof paradox is explaining why bare statistical evidence seems insufficient for legal verdicts; must further epistemic conditions be satisfied?

This paper remains neutral about the epistemic justification of the target verdicts themselves. That is, I do not judge whether the Taxi One and Taxi Two legal verdicts (and, later, the lottery, iPhone, and prisoner vignette judgements) are justified. I focus on the similarities and overlooked differences amongst the pairs of vignettes, what the paradoxes are, and whether Pritchard's safety account vindicates the results he claims.

Pritchard argues that his safety condition explains the epistemological flaw of the verdict based on bare statistical evidence.[4] Pritchard argues that typically an outcome's being statistically unlikely and its being a remote possibility go together. Something's being probabilistically unlikely reflects the fact that its obtaining requires the world to be somewhat different from how it is. It is unlikely that Vladimir Putin moves to Zimbabwe next year, and this unlikelihood

[3] See E. M. Morgan 'The Law of Evidence, 1941–45' *Harvard Law Review* 59 (1946) 481–576; Cohen op. cit. note 2, Thomson op. cit. note 2.

[4] Duncan Pritchard 'Epistemic Risk' *Journal of Philosophy* 113 (2016) 550–571; Duncan Pritchard 'Legal Risk, Legal Evidence and the Arithmetic of Criminal Justice' *Jurisprudence* 9 (2018) 108–119. See also Michael S. Pardo 'Safety vs. Sensitivity: Possible Worlds and the Law of Evidence' 24 *Legal Theory* (2018) 50–75.

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

reflects the fact that a lot would have to be different for this to happen. But sometimes probabilistic unlikelihood and modal distance come apart: the statistical improbability of the outcome does not reflect its modal distance; the outcome could obtain without a significant change in the world. This is illustrated by lotteries: it is unlikely that my numbers win this week, but it is not modally distant. Not much would have to change in the world for this to happen. Pritchard argues that the distinction between mere statistical unlikelihood and being a remote possibility explains why playing the lottery is attractive, and even somewhat rational. The slogan 'it could be you' speaks to the modal nearness of winning.

Pritchard holds that low statistical probability of error does not suffice for an appropriate legal finding of culpability. In addition, the low statistical probability must reflect that the error is a modally distant possibility. The case against the Red Taxi company based on bare statistical evidence provides the former condition, but it fails to provide the latter.

In Pritchard's terminology, the Red Taxi verdict based on bare statistical evidence is unsafe. The verdict could too easily be wrong. Pritchard understands the easy possibility of error using a Lewisian possible worlds framework:[5]

> **Safety.** S's belief is safe if and only if in most nearby possible worlds in which S continues to form her belief about the target proposition in the same way as in the actual world, and in all very close nearby possible worlds in which S continues to form

[5] Duncan Pritchard *Epistemic Luck* (Oxford: Oxford University Press, 2005); 'Anti-Luck Epistemology' *Synthese* 158 (2007) 277–297; 'Safety-Based Epistemology: Whither Now?' *Journal of Philosophical Research* 34 (2009) 33–45; 'Anti-Luck Virtue Epistemology' *Journal of Philosophy* 109 (2012) 247–279. Although Pritchard's formulation divides nearby worlds into two discrete classes – nearby possible worlds and very close nearby possible worlds – this is best understood as a continuum: closer worlds are more significant for assessments of safety (2012: 255). For criticisms of Pritchard's safety-based approach to knowledge, see Juan Comesaña 'Unsafe Knowledge' *Synthese* 146 (2005) 395–404; John Greco 'Worries about Pritchard's Safety' *Synthese* 158 (2007) 299–302; Avram Hiller and Ram Neta 'Safety and Epistemic Luck' *Synthese* 158 (2007) 303–313; Pritchard (2009) op. cit. note 5; Lisa Miracchi 'Competence to Know' *Philosophical Studies* 172 (2015) 29–56; Georgi Gardiner 'Safety's Swamp: Against the Value of Modal Stability' *American Philosophical Quarterly* 54 (2017) 119–129; Maura Priest 'Why Anti-Luck Virtue Epistemology has No Luck with Closure' *Logos and Episteme* 8 (2017) 493–515.

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

her belief about the target proposition in the same way as in the actual world, the belief is true.

To evaluate whether S's true verdict is safe, consider the nearest possible worlds in which S believes p on the same basis. If there are nearby worlds where S believes p on that basis, but p does not obtain, the belief is unsafe. More precisely, the belief is unsafe if there are any very close nearby worlds – or a significant proportion of more distant nearby worlds – in which S believes p (on the same basis), yet not p.

Pritchard employs his safety condition to diagnose the epistemic inadequacy of the Taxi One verdict: The verdict is unsafe.

3. The Lottery Paradox and the Profiling Paradox

Pritchard applies the same apparatus to Harman's lottery paradox.[6] The lottery paradox is, like the proof paradox, constructed by contrasting two vignettes:

Lottery One. Lottie purchased a ticket in an enormous national lottery. Millions of tickets are sold. The draw has just occurred, but Lottie has not inquired about the result. Reflecting on the odds, she reasons that her ticket has not won. 'My ticket hasn't won', she asserts, as she discards it.

Lottery Two. Cammie purchased a ticket in her community lottery, which has 50 tickets. She reads the results in her local newspaper, and the number does not match her ticket. 'My ticket hasn't won', she asserts, as she discards it.

The structure of this pair of cases resembles the proof paradox cases.[7] In the first case the judgement is very likely correct, but the bare statistical evidence available does not seem to legitimate the relevant assertion and action. In the second case, non-statistical, 'individualised' evidence is

[6] Gilbert Harman 'Knowledge, Inference, and Explanation' *American Philosophical Quarterly* 5 (1968) 164–173

[7] Redmayne op. cit. note 2; Lara Buchak 'Belief, Credence, and Norms' *Philosophical Studies* 169 (2014) 285–311; Duncan Pritchard, 'Risk' *Metaphilosophy* (2015); Pritchard op. cit. note 4, and others explicitly state that the taxi examples closely resemble, or are identical to, lottery examples. As I argue in section five, the taxi vignettes exhibit significant epistemological differences from lottery vignettes. The prisoner version of the proof paradox more closely resembles the lottery vignettes, although there are nevertheless differences.

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

available, and Cammie's actions and assertions seem legitimate. Given that the lottery in the second case is significantly smaller, Cammie's ticket is considerably more likely to win than Lottie's. Even given the additional evidence – the local newspaper – the overall probability of the target fact, 'My ticket hasn't won', given the total evidence might well be lower for Cammie than for Lottie, yet action and assertion nonetheless seem more appropriate given Cammie's evidence. Arguably Cammie knows that her ticket has not won, and Lottie lacks knowledge. Yet plausibly Lottie's belief is more likely true, given the available evidence. This is Harman's lottery paradox. The challenge raised by the lottery paradox is explaining the epistemic normativity exemplified by the two cases; why does bare statistical evidence seem insufficient for knowledge in Lottery One?

Pritchard's account offers the same explanation for Harman's lottery paradox and the proof paradox: the judgement formed in 'Case One' – the verdict based on bare statistical evidence – is unsafe. The judgement based on individualised, non-statistical evidence, by contrast, is safe. Pritchard writes, 'In effect, what we have [with the taxi vignettes] is a legal lottery case (albeit where the odds are not so massively stacked against the target event)'.[8] On Pritchard's view, it makes no epistemic difference whether the judgements are about people who have agency or moral interests. The fact that the verdict concerns persons is epistemically inert. Since it makes no epistemological difference to Pritchard's analysis, the city's taxi drivers in the vignettes could be replaced by objects lacking agency and moral significance, like cascading Galton Board beads. Pritchard's view thereby understands the proof paradox as resolvable through the apparatus of standard, orthodox epistemology. There is no appeal to moral or social constraints, or to distinctive features of legal systems.

This observation raises questions: Does Pritchard err in treating the proof paradox and Harman's lottery paradox in the same way? Do the vignette pairs exemplify the same central epistemological features? To address these questions, consider a pair of vignettes introduced by Buchak.[9]

> **iPhone One.** Fiona's iPhone is stolen. She knows that only two people, Jake and Barbara, entered the building. And she knows that nine out of ten iPhone thefts are committed by men. On this evidence Fiona believes that Jake stole the phone. Fiona's belief is true.

[8] Pritchard, op. cit. note 4, 118.
[9] Op. cit. note 7.

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

> **iPhone Two.** Fiona's iPhone is stolen. She knows that only two people, Jake and Barbara, entered the building. Men and women steal equal proportions of iPhones. An eyewitness from a considerable distance says she saw Jake steal the iPhone. On this evidence Fiona believes that Jake stole the phone. Fiona's belief is true.

Buchak treats this pair of cases as having an identical structure to the original taxi proof paradox vignettes. She writes, 'Here is another case, with the same form, that seems to prompt the same intuitions'.[10]

Buchak introduces the iPhone cases to further illustrate the proof paradox, but it will be helpful to introduce the term 'profiling paradox'. The 'profiling paradox' asks why Fiona's (true) iPhone One belief (based on bare statistical evidence) is unjustified but her iPhone Two belief (based on individualised evidence) could be justified, even if the former is more statistically likely given the evidence. Here 'profiling' refers to forming opinions about an individual on the basis of statistics about their social group. The challenge raised by the profiling paradox, as exemplified by the iPhone cases, is explaining the epistemic error, if any, of judging an individual on the basis of base rates about his social group, even in cases where plausibly the demographic evidence renders the relevant claim likely.

On Buchak's view the crucial distinction is whether the evidence supports belief or merely very high credence. Bare statistical evidence can support high credence, Buchak holds, but it does not support belief. And, Buchak continues, belief and credence play different normative roles. Credence governs the domain of expected utility theory, but bare credence – even very high credence – cannot support reactive attitudes, blame, punishment, incarceration, and so on. As Buchak observes, you cannot half-blame someone, on a middling credence that they wronged you.[11] You instead need full belief that they acted wrongly.

Buchak presents the iPhone vignettes as a simple two-person proof paradox, akin to the taxi vignettes. Pritchard concurs; he also treats the iPhone vignettes as epistemologically isomorphic to the taxi vignettes. This is mistaken. To see why, apply Pritchard's resolution of the proof paradox to iPhone One: Drawing only on base rates, Fiona believes Jake stole the iPhone. Is her belief safe?

Whether Fiona's belief is safe depends on modal features of Jake's theft. In what follows I argue it is not mere random chance whether Jake or Barbara stole the iPhone. It is not a random lottery and so, at

[10] Op. cit. note 7, 292.
[11] Op. cit. note 7, 299.

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

least given Pritchard's account, cannot be simply treated like one for epistemological purposes. The case involves profiling, and profiling differs epistemically from mere statistical lottery chances.

To understand the modal features of Jake's theft, we must examine what underwrites the (stipulated) social statistic that nine out of ten iPhones thefts are committed by men. I turn now to stipulative epidemiology to illuminate what can underlie demographic differences in disease incidence. As I illustrate, underlying causal differences give rise to modal differences. This, I argue, casts doubt on Pritchard's safety-based diagnosis of the inadequacy of bare statistical evidence for outright verdicts.

4. Stipulative Epidemiology: Underlying Causes and Modal Structure

Suppose there are three common diseases, A, B and C. Approximately half the total population have the disease, and half the total population is male. For simplicity, the diseases begin in youth and are chronic. These diseases are fictional.[12]

Statistic One. 90% of people with disease A are male
Statistic Two. 90% of people with disease B are male
Statistic Three. 90% of people with disease C are male

In a superficial sense, there is no statistical difference among facts one to three. But there is a difference in what underlies the statistic in each case.

With disease A, there is a scale measuring how poorly an individual regulates sugar. Disease A marks a threshold on that scale. Many people regulate sugar poorly to some degree; if sugar regulation lies beyond the threshold, the person has disease A. Although disease A is fictional, it resembles type-2 diabetes in some respects.[13] Statistic One arises because although males and females regulate

[12] In reality, most prevalent diseases are multi-factorial and cannot be modelled simply; epidemiologically simple diseases are rare diseases. The American Heart Association reports cardiovascular deaths represent 30% of deaths globally. Nearly 50% African-American adults have cardiovascular disease. Cancer Research UK estimates the lifetime cancer risk is 50% for males, and 45% for females. I am grateful to Jelena Aleksic, Ilya Chevyrev, and Zoltán Molnár for helpful discussions.

[13] The CDC reports >9.3% of the US have diabetes, and >33% US adults have prediabetes, where blood sugar levels are higher than normal but not high enough to classify as diabetes. 15–30% of prediabetics

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

sugar poorly to some degree, males tend to be considerably worse. And so more males than females qualify as having disease A. Males generally tend to be worse than females at regulating sugar, whether or not they have the disease.

With disease B, the condition is binary, and there is no underlying scale. Individuals either have the medical problem or they do not, much like person either has (for illustration) colon cancer or they do not. This contrasts with disease A, where diagnosis marks a place on a spectrum of poor sugar regulation and many people who do not qualify as having the disease nonetheless have the medical problem to some degree.

Statistic Two arises because male anatomy is more susceptible to disease B than female anatomy. And so more males than females contract the disease. Even those males who have not yet contracted disease B might well do so. Their body shape and body chemistry are risk factors. Female bodies tend to be more protected against the disease. It might be helpful to think of something like colorectal cancer or groin hernia, as real-world analogues for disease B.[14] If you have a male friend and a female friend, and neither has colon cancer, you would worry more about the male friend contracting it; it is a closer possibility. You would sooner want him to get checked out.

Disease C is also binary and does not mark a threshold on an underlying scale, and more males have it than females. But those males who do not contract disease C are no more prone to disease C than females without disease C. To better illustrate this modal structure, I will put some stipulated flesh on the bone. The underlying causation that underwrites Statistic Three is – let's say – there is a habit H, chewing tobacco. Disease C is mouth cancer. Far more males than females develop habit H. Habit H correlates extremely strongly with disease C: there are almost no other causes of mouth cancer and habit H

develop diabetes within five years. Black and Hispanic adults are about twice as likely to develop diabetes as non-Hispanic white adults.

[14] Approximately 27% of males and 3% of females develop a groin hernia; RJ Fitzgibbons Jr and RA Forse 'Clinical Practice. Groin Hernias in Adults' *The New England Journal of Medicine* 372 (2015) 756–63. The male-to-female ratio for colorectal cancer is 1.35 and for bladder cancer is 4.0; M. Tefvik Dorak and Ebru Karpuzoglu 'Gender Differences in Cancer Susceptibility: An Inadequately Addressed Issue' *Frontiers in Genetics* 3 (2012) 268–279. Prostate and Skene gland cancer might offer a potential real-world analogue for disease B and – if the sexes were reversed – so might breast cancer.

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

often leads to mouth cancer. But males without habit H are no more likely than females without habit H to contract disease C.[15]

Even though the statistical gender distribution is the same, as articulated by facts one to three, the modal patterns are different. The common statistics disguise the modal patterns that might make an epistemological difference; they certainly make a difference according to Pritchard's modal epistemology.

Suppose S forms a true belief that Manuel has disease X based on the fact that Manuel is a man, the bare statistical facts one to three, and the disease's high prevalence. (Since half the population are male and approximately half the population contracts disease X, approximately 90% of males contract disease X.) Is S's true belief safe?

If the disease in question is disease C, plausibly S's belief is safe. The nearest world in which Manuel does not contract disease C is a non-habit world. And plausibly if Manuel has habit H, then non-habit worlds are relatively distant.

If the belief concerns disease B, by contrast, Manuel has the at-risk body shape. But plausibly he might easily have not contracted the disease. Some diseases, such as certain cancers, have a randomness, or 'luck of the draw'. Plausibly there could be identical twins, with the same body shape, environmental conditions, diet, and exercise habits, yet one contracts cancer and the other does not. There is modal closeness – happenstance – or bad luck about contracting some cancers. This modal nearness can threaten the safety of S's belief about disease B, at least relative to beliefs about disease C.

For disease A, we need more information to evaluate the belief's safety. Perhaps Manuel is borderline for the disease threshold. If so, in many nearby worlds Manuel lacks disease A, and S's true belief is unsafe. On the other hand, Manuel's sugar regulation could be very poor indeed. If so, in all nearby worlds Manuel has disease A and S's belief is safe.

Suppose Manuel does not suffer from the target disease. Let's now briefly examine the modal profile of S's false belief. If Manuel lacks disease C, then very likely Manuel lacks habit H, and plausibly he is

[15] Males smoke more (habit H) and contract corresponding diseases (disease C) at a higher rate; the male-to-female lung and bronchus cancer ratio is 1.52 (Dorak and Karpuzoglu op. cit. note 14). The male-to-female tuberculosis ratio approximates 2:1. In some nations smoking might explain up to one third of this sex difference; Shepherd Nhamoyebonde and Alasdair Leslie 'Biological Differences Between the Sexes and Susceptibility to Tuberculosis' *The Journal of Infectious Diseases* 209 (2014) s100–s106.

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

no more at risk than an arbitrary non-habit woman. He only contracts disease C at distant (habit H) worlds. By stipulation Manuel lacks disease B but, by contrast to disease C, he is still at risk. Plausibly in nearby worlds he has disease B; not much would need to be different for Manuel to contract disease B. With disease A, as above, it depends on where along the sugar-regulation scale Manuel actually is. If he is close to the threshold, Manuel will have the disease in nearby worlds.

Since Buchak's iPhone vignettes concern crime, not disease, we need to consider the modal structure underlying the demographic crime statistic that Buchak stipulates. The modal pattern depends on what underlies the gendered crime distribution. Theorists can stipulate, speculate, or estimate the causes and modal patterns of (actual or fictional) crime distribution. As the foray into stipulative epidemiology indicates, if there is a large gender difference – like the one Buchak stipulates in iPhone One – it is unlikely to be a 'pure chance' random lottery.

It is difficult to imagine diseases with a substantial (90:10) gender distribution in which incidence is nonetheless a pure, random lottery, characterised by Pritchardian flukiness and modal closeness. This difficulty is because something *causes* the gender difference; and this cause means contracting the disease is probably not truly random, like a lottery.

We can stipulate a disease D: Radiation travels across space and some passes through Earth. If it hits an M-molecule, the radiation causes cancer. Having M-molecules in your body is like participating in a radiation lottery. The gender difference is caused by the fact that males have significantly more M-molecules than females.[16] Males have more tickets in the radiation lottery, where M-molecules are lottery tickets, and 'winning' is cancer. Disease D – which does not resemble any disease I know of – illustrates the possibility of random lotteries that exhibit gendered differences. But these are unusual, abnormal cases. Normally if there is a gendered difference, the details of the causal structure mean that contracting the disease is not best understood as a random lottery.

5. Probing the Profiling, Proof, and Lottery Paradoxes

The key point is that the taxi vignettes that exemplify the proof paradox might be plausibly construed as epistemologically similar

[16] Testosterone or cremaster and inguinal canal cells might be real-world analogues of M-molecules. On average, adult male testosterone levels are approximately eight times higher than adult female levels.

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

to the lottery paradox vignettes (perhaps illegitimately; I return to this).[17] But the iPhone vignettes cannot. Buchak's iPhone One involves profiling, and this is essential to the case. And so the question arises for iPhone One: What underwrites the demographic difference?

Buchak treats the three pairs of vignettes alike: Lottie and Cammie of the lottery paradox, the two taxi vignettes comprising the proof paradox, and the two iPhone vignettes comprising the profiling paradox. In the first vignette of each pair, S should not form a verdict, on Buchak's view, because S only possesses bare statistical evidence, and accordingly only credence is warranted.

But on Pritchard's view a chasm appears. If Lottie believes her ticket did not win based on bare statistical evidence – the size of the lottery – her verdict has a defect. It is unsafe. On the available evidence, she forms a false belief in modally nearby worlds. In the taxi case, Pritchard argues, the verdict based purely on market share evidence is unsafe.

But when we consider the iPhone vignettes, we must confront the fact that if Jake stole the phone, he stole the phone in nearby worlds – that is, in all very close nearby worlds, and almost all slightly further nearby worlds. Theft is not typically random. It might be opportunistic, unplanned, and so on. But it is not a random lottery between whether Jake or Barbara stole the phone. It is not simply a 90% chance that Jake stole the phone and a 10% chance that Barbara did, where it could easily have been Barbara instead; the iPhone case differs modally from a lottery case. Unlike with a lottery vignette, a lot would need to change for Barbara, rather than Jake, to be culpable. Fiona's true belief that Jake stole the iPhone is safe.

If the iPhone case is not compelling, consider the modal profile of sexual assault, as a variation of the iPhone case. Suppose a sexual assault occurs and the perpetrator was either Jake or Barbara. Based just on the same kind of evidence – the gender distribution of the crime – plausibly Fiona's belief that Jake was the culprit is safe, or at least significantly safer than the corresponding Lottery One belief.[18] (I do not claim Fiona's belief is justified, only that on

[17] I argue that the prisoner vignettes, introduced below, are more similar to the Harman lottery vignettes and Kyburg lottery paradox than the taxi vignettes are.

[18] Note, however, that Fiona's belief, based on bare statistical evidence, that 'Jake committed the crime' is insensitive. If the claim were false, S would nonetheless believe it. If intuitively Fiona's judgement is flawed, this result favours the sensitivity condition as an explanation for the inadequacy of bare statistical evidence.

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

Pritchard's view, the belief is safe. Recall this essay remains agnostic about the epistemic justification of the verdicts themselves.)

In light of this, could Pritchard revise the safety condition so that Fiona's belief about Jake counts as unsafe? I am doubtful. Given that a sexual assault occurred, the belief that Jake is guilty is extremely modally stable. If Fiona's belief about Jake does not qualify as safe, then – given how demanding this revised safety condition must be – few beliefs qualify as safe. To see why, suppose Pritchard concedes that on this occasion Jake, not Barbara, is the culprit in all nearby worlds; Jake's guilt is modally stable. But, Pritchard continues, Fiona's belief that Jake committed the crime is nonetheless unsafe because, *were* Fiona to form beliefs on this kind of evidence on many *other* occasions – about other men – the method could lead to a false belief in a different (albeit modally distant) case. Since some of Fiona's (non-actual) beliefs, formed on a similar basis, could be false on different (non-actual) occasions, her actual belief about Jake is unsafe. Pritchard writes,[19]

> Imagine… you're a judge faced with a series of cases [with this kind of inculpatory evidence]. You know that in 90% of cases of this kind the man is guilty. Accordingly, you decide that no further evidence is necessary and so set about convicting, one by one. The evidence is excellent, statistically speaking (make it 99% if you're not convinced of this), but the basis for belief is clearly unsafe, as the judge could very easily convict an innocent man via this process.

In response, note Fiona is not a judge; she simply forms a belief about Jake. And her resulting belief is true in all nearby worlds. To emphasise: I do not endorse Fiona's belief, I merely claim it satisfies Pritchard's classic safety condition. Note also that Pritchard's proposed response to the sexual assault example more resembles sensitivity than safety. It asks: Were Fiona to form beliefs in other (possibly distant) cases, and in those cases the proposition were false, would her belief track the truth in those counterfactual not-p cases?

I concede that if satisfying this more demanding condition is necessary for a belief to qualify as safe, then plausibly Fiona's belief about Jake is unsafe; Pritchard thereby secures the right result about the vignette. But this revised safety condition is implausibly demanding: In the vignette, Fiona forms a belief about Jake. Her belief is true in all nearby worlds. If Fiona's actual belief is nonetheless unsafe

[19] Personal communication, April 2018. I am extremely grateful to Duncan Pritchard for many helpful conversations about these topics.

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

because a different (counterfactual, modally distant) belief on similar evidence in a different (counterfactual, modally distant) case is false, then safety is a very demanding standard. Fiona actually only forms a belief about Jake. And so the possible worlds in which Fiona forms similar beliefs about other men on similar evidence, and her belief is false, are distant worlds. The falsity of those modally distant beliefs, this revised safety account holds, renders Fiona's actual belief unsafe. (Recall that Fiona's actual belief – the one about Jake – is true in all nearby worlds.) Many beliefs do not qualify as safe, according to this more demanding account of safety.[20]

Thus it seems that Pritchard's safety account *may* explain why the verdicts based on bare statistical evidence in Taxi One and Lottery One are unwarranted, but cannot explain any fault with the demographic profiling in iPhone One. It cannot identify an epistemic fault with (at least some kinds of) profiling. Perhaps Pritchard is correct that Case One verdicts are unjustified because S could 'too easily be wrong', but Pritchard's safety condition does not capture this notion of easy error possibility.

At the outset I introduced the proof paradox by contrasting two vignettes, Taxi One and Taxi Two. In Taxi One, purely statistical evidence did not legitimate a verdict of legal culpability; but in Taxi Two, individualised, non-statistical evidence could suffice, even though the overall probability of culpability given the evidence was lower. This contrast is the heart of the proof paradox. The proof paradox – and the intuitive inadequacy of bare statistical evidence for legal judgement – is often instead introduced by contrasting the following two vignettes, adapted from Nesson.[21]

> **Prisoner One.** One hundred prisoners exercise in the yard. Security footage reveals that ninety-nine prisoners together attack a guard; one prisoner refuses to participate. Prison officials decide that since for each prisoner it is 99% likely they are guilty, they have adequate evidence to successfully prosecute individual

[20] An example: I ask my friend her grandmother's name. She replies truly and would not speak falsely in nearby worlds. Intuitively my resulting actual belief is safe and known; not easily could it be wrong. But this belief violates the more demanding proposed standard: Were I to instead ask many people, on many occasions, eventually likely someone would reply falsely. They would be mistaken or deceptive. Thus the proposed Pritchardian reply impugns many ordinary beliefs.

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

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

prisoners for assault. They charge Ryan with assault. A guilty verdict is returned.

Prisoner Two. One hundred prisoners exercise in the yard. Ten prisoners attack a guard. A second security guard, standing at a substantial distance, believes that he saw that Ryan was one of the attackers. Prison officials charge Ryan with assault. A guilty verdict is returned.

This version of the proof paradox vignettes also shares key relevant features with the vignettes of Harman's lottery paradox and the profiling paradox. Plausibly in Lottery One and Prisoner One the bare statistical evidence licenses a high credence in the target claim (perhaps 0.99). But something is amiss with establishing culpability on this evidence alone. In Lottery Two and Prisoner Two non-statistical, individualised evidence is available. The overall probability of the disputed fact given the evidence might well be lower, yet the relevant burden of proof – beyond reasonable doubt – could well be satisfied on this basis.

Other than illustrating different legal standards – preponderance of evidence and beyond reasonable doubt – the taxi and prisoner examples are typically treated as epistemically alike. They are presented as exemplifying the same epistemic features, and so resolvable by positing the same epistemic requirements on legal verdicts. Mike Redmayne, for example, writes,

> Lottery is striking for its similarity to Prisoners… the similarities between Lottery and Prisoners suggest that verdicts require knowledge… It is plausible that whatever prevents a liability verdict in [Prisoner One] also prevents a liability verdict in [Taxi One]. If Prisoners is explained by a knowledge requirement for proof, then [the taxi example] is too.

Redmayne articulates explicitly two widely-held implicit assumptions:[22] Firstly, the lottery vignettes are epistemically akin to the prisoner vignettes. Secondly, the taxi vignettes share this similarity. The epistemic defect exhibited by Taxi One is the same as the epistemic

[22] Smith op. cit. note 2; Pardo op. cit. note 2; Ho Hock Lai *A Philosophy of Evidence Law* (Oxford: Oxford University Press, 2008); Redmayne op. cit. note 2; Pritchard op. cit. note 4; Buchak op. cit. note 7. Note a vignette can exhibit multiple epistemic errors. This means that Case One judgements could exemplify some shared epistemic faults, such as outright judgement on bare statistical evidence, whilst differing in other epistemic properties, such as whether the evidence is rendered inadequate by high error costs.

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

defect of Prisoner One and Lottery One. Pritchard, Buchak, and others hold this epistemic defect is also the same as that exhibited in iPhone One. This assumption warrants scrutiny.

These paired vignettes have similarities. But they also exhibit overlooked differences. The lottery and prisoner vignettes share features with the vignette comprising Kyburg's lottery paradox and Makinson's preface paradox.[23] Namely, if applied to each ticket or prisoner in the example, a false verdict is inevitable. This is guaranteed by the example's structure. This contrasts with the taxi, iPhone, and sexual assault vignettes. In these cases, repetition does not guarantee an eventual false belief; a judge could condemn individuals on the available base rate evidence for decades, and never rule falsely. This overlooked difference bears on responses to the proof paradox. Martin Smith, for example, proposes that the reason bare statistical evidence does not suffice for conviction, as illustrated by the prisoner proof paradox, is that convicting one person on that evidence commits us to convicting all – for fairness – but '[b]y convicting all 100 individuals, however, we are now guaranteed to have convicted an innocent – we have more convictions than we do crimes… No legal system that respects a right against wrongful conviction can tolerate an oversupply of guilty verdicts – more guilty verdicts than there are crimes'.[24] Smith overlooks that his 'surfeit of convictions' strategy does not readily extend to the taxi proof paradox.[25]

Secondly, in the lottery and prisoner vignettes, there is no additional evidence to draw on, such as evidence that distinguishes any particular ticket or prisoner, or circumscribes epistemically significant reference classes within the total set of tickets or prisoners. There is very little – perhaps nothing – to infer about individuals beyond what is stated explicitly. The taxi, sexual assault, and iPhone cases differ: they essentially rely on inferences from background base rates to the individual on the scene. Group-to-individual inference, and selection of the relevant reference class, play a crucial

[23] Henry E. Kyburg *Probability and the Logic of Rational Belief* (Wesleyan University Press, 1961); D. C. Makinson 'The Paradox of the Preface' *Analysis* 25 (1965) 205–207.

[24] Smith op. cit. note 2.

[25] See also discussion of 'capped convictions' (Littlejohn, op. cit. note 2). As Smith (op. cit. note 2) notes, in the lottery and preface paradox vignettes once you endorse each individual claim, there is a further question, namely whether to conjoin individual verdicts and endorse the conjunction. The prisoner and taxi cases lack this option. Once you convict each individual, there is no isomorphic further question of whether to additionally convict the conjunction.

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

role.[26] Plausibly, then, the taxi, sexual assault, and iPhone vignettes crucially concern the epistemology of profiling, whereas the prisoner and lottery vignettes do not.

To see this distinction clearly, consider the reference class problem. One could easily use the taxi, sexual assault, or iPhone vignettes to illustrate the reference class problem: Should we reason from the proportion of total taxis in the city, or in the vicinity, or on duty that evening, or miles driven, or collision histories, or 'hit-and-run' records…? Perhaps the Red Taxi company employs the safest drivers. The taxi vignette is run through with the reference class problem! It is more challenging to employ the lottery or prisoner vignettes to illustrate the reference class problem. This is largely because the lottery and prisoner vignettes, as stipulated, lack significant choices about which group can support a profiling inference.[27] There is a group-to-individual inference in the lottery and prison cases, but the *selection* of the relevant social group has no significant or interesting role in the inference. This contrasts with the iPhone and Taxi cases, where selection of the relevant social group is integral to the case.

Thus, there are rich, unappreciated epistemic differences amongst the central cases used to motivate the proof paradox – the prisoner and taxi cases – and between those cases and the lottery case, on the one hand, and the iPhone and sexual assault examples of the profiling paradox, on the other hand.

These observations suggest we might turn an even more critical eye to Pritchard's account. I have emphasised the significance of what underwrites a demographic statistic – I illustrated how identical statistical facts might obscure important modal differences – and I suggested ways that behaviour and disease do not resemble a random lottery case. I thereby argued that Fiona's belief in iPhone One (and the sexual assault variation) is safe, and so Pritchard's account cannot explain the epistemic flaw of at least some kinds of profiling.

[26] Several interlocutors have suggested the crime distribution articulated in iPhone One says *nothing* about the probabilities concerning individuals in the building, but in Prisoner One the base rates do reveal something about the individuals. I demur, but this suggestion gestures at this difference between the iPhone and prisoner cases.

[27] Realistic prison populations exhibit subgroups, granted, but these reference classes are not essential to the case. By contrast, we cannot describe the taxi vignette without substantial assumptions about the relevant reference class. Thanks to Andy Mueller, Jon Garthoff, and participants in Sherri Roush's epistemology seminar at UCLA for helpful discussion.

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

But this criticism may cut deeper. This reasoning suggests it is mistaken to treat the *taxi* proof paradox vignette as epistemically akin to a lottery, as Pritchard does. Perhaps the true verdicts formed in taxi cases *are* safe; drivers who 'hit and run' would do so in nearby possibilities. And drivers who do not 'hit and run' would not do so in nearby possibilities. The Taxi One verdict, although based on bare statistical evidence, is safe.[28] Similarly, perhaps it is mistaken to treat Prisoner One as epistemically similar to a lottery, as Pritchard does. Plausibly true verdicts formed in that case are safe; individuals who riot would riot in nearby worlds. The prisoner who refuses to riot would similarly refuse in nearby worlds. This suggests that the only epistemic flaw safety theorists can appeal to is the vanilla, ordinary flaw that one of the verdicts (if repeated for each prisoner) is false. Safety drops out of the picture.[29]

6. On Seeking a Unified Explanation of the Paradoxes

I examined epistemic features of five pairs of vignettes – Lottery, Prisoner, Taxi, and iPhone, with its sexual assault variation. This discussion raises two questions. Firstly, is the verdict warranted in each of the vignettes? Received wisdom holds the verdict described in Case

[28] Granted, were a judge to often rule against taxi drivers on this evidence, she might eventually rule falsely. But this does not mean the original Taxi One verdict violates Pritchard's classic safety condition; it simply means if the judge continually repeated her method, one day she might judge falsely. As discussed above, many beliefs are safe, according to Pritchard's classic analysis, even though the method can potentially lead to false belief. This includes, for example, belief about a friend's grandparent's name, formed on the basis of asking that friend. Were you to ask many friends, you might one day form a similar belief falsely via a similar method. But this does not impugn the safety of your actual belief, which is modally secure. Put another way, if beliefs are unsafe merely because one day the method (might) result in a false belief, safety is too demanding.

[29] A further objection to Pritchard's treatment of the proof and profiling paradoxes is available to advocates of moral encroachment. The objection proceeds: Some vignette judgements are morally weighty, and so by treating the vignettes alike Pritchard thereby overlooks important aspects of epistemic and moral normativity. I describe moral encroachment in section six. For further criticisms of a safety-based approach to explaining the proof paradox, see Georgi Gardiner 'Legal Evidence and Knowledge' in Maria Lasonen-Aarnio and Clayton Littlejohn (eds.) *The Routledge Handbook of the Philosophy of Evidence* (Routledge, forthcoming).

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

One in each pair – the verdict based on bare statistical evidence – is unwarranted, or is at least less warranted than the corresponding Case Two verdict. This paper remains agnostic about the first question, and instead charts overlooked epistemic contours of the vignettes. The second question asks whether a unified account of the epistemology of the vignettes is tenable or even desirable. Can the same epistemological tools explain whether or not the verdicts are warranted in each vignette?

These two questions come apart. The same epistemological features might perform the explanatory work, where that work is explaining why some Case One verdicts are warranted, whilst others are not. Conversely, perhaps different epistemological posits are operative for different cases, yet the answers to the first question are the same across the pairs of cases. To illustrate, consider the sensitivity condition:

> **Sensitivity.** S's true belief that p is sensitive iff in the nearest possible worlds in which p is not true, S does not believe p.

Plausibly the sensitivity condition diagnoses that all Case One verdicts are unwarranted, and for the same reason: The resulting judgement is insensitive.[30] Given the evidence – base rates, crime distributions, lottery size, and so on – in the nearest worlds where not p, S nonetheless believes p. A unified explanation diagnoses homogenous judgements about each Case One verdict. Suppose instead a theorist holds that verdicts based on bare statistical evidence are unwarranted in Lottery One, but warranted in iPhone One. They might marshal a safety condition – a unified explanation – to explain these heterogeneous judgements: The Lottery One verdict is unsafe, but iPhone One is safe, even though both are based on bare statistical evidence.

Orthodoxy holds that if a single epistemological posit explains intuitive judgements about all the vignettes, this qualifies as a virtue of the account. If one posit – a safety condition, say – can explain a range of epistemic phenomena and intuitions, the posit is thus deemed more powerful and fruitful. Orthodoxy thereby encourages theorists to apply the same epistemological posit to diagnose the answers to the first question. A recent wave of epistemological theorising argues,

[30] Cf. Enoch, Spectre, and Fisher op. cit. note 2; David Enoch and Talia Fisher 'Sense And "Sensitivity": Epistemic and Instrumental Approaches to Statistical Evidence' *Stanford Law Review* 67 (2013) 557–611; David Enoch and Levi Spectre 'Statistical Resentment' (ms). See also Sherrilyn Roush *Tracking Truth* (Oxford: Oxford University Press, 2007).

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

however, there is something deeply mistaken about treating the cases alike. Moral encroachment holds that moral features of a judgement can affect its epistemic justification.[31] There are several varieties of moral encroachment, but most treat the various vignettes as epistemologically heterogenous.

Rima Basu and Mark Schroeder would likely argue Prisoner One, Taxi One, and iPhone One exhibit high moral stakes, and so more evidence is needed to justify the verdicts.[32] This explanation is compatible with Lottery One verdicts being warranted, since they are morally neutral. (Some theorists might aver that iPhone One and the sexual assault variation do not exhibit high stakes, because Jake is not a member of a marginalised group.)[33] Sarah Moss suggests that when forming judgements about individual people, the possibility that the individual differs from others in their reference class is a salient relevant alternative, and so Case One verdicts are unwarranted if the judgements concern individuals.[34] Whether Case One verdicts that are not about people, such as Lottery One, are warranted is explained by different epistemological apparatus.

These moral encroachment views apply heterogenous epistemological explanations to the various cases. They hold that treating the cases alike overlooks important aspects of epistemic and moral normativity. On Pritchard's view – indeed, according to orthodoxy – it is a virtue of his account if his safety posit explains the epistemological contours of all the cases. Moral encroachment, by contrast,

[31] Michael Pace 'The Epistemic Value of Moral Considerations: Justification, Moral Encroachment, and James' 'Will to Believe'' *Noûs* 45 (2011) 239–268; James Fritz 'From Pragmatic Encroachment to Moral Encroachment' *Pacific Philosophical Quarterly* (2017); Sarah Moss *Probabilistic Knowledge* (Oxford: Oxford University Press, 2018); Rima Basu 'Radical Moral Encroachment: The Moral Stakes of Racist Beliefs' *Philosophical Issues 29* (2019) 9–23; Bolinger op. cit. note 2; Renee Bolinger 'Varieties of Moral Encroachment' (ms). For criticisms of moral encroachment, see Georgi Gardiner 'Evidentialism and Moral Encroachment' in Kevin McCain (ed.) *Believing in Accordance with the Evidence* (Springer, 2018) 169–195.

[32] Rima Basu and Mark Schroeder 'Doxastic Wronging' in Brian Kim and Matthew McGrath (eds.) *Pragmatic Encroachment in Epistemology* (Routledge, 2019) 181–205.

[33] Cf. Rima Basu *Beliefs That Wrong*, doctoral dissertation, (University of Southern California 2018).

[34] Op. cit. note 31. For criticism of Moss's moral encroachment view, see Georgi Gardiner 'Relevance and Risk: How the Relevant Alternatives Framework Models the Epistemology of Risk' *Synthese* (forthcoming).

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

holds that treating vignettes about crimes as epistemologically akin to vignettes about lotteries and Galton Boards is mistaken; epistemic normativity ought to be sensitive to moral features. On this view, the paradoxes constituted by morally loaded vignettes exemplify very different epistemological phenomena from those constituted by morally neutral vignettes. A unified account of these paradoxes, is neither tenable nor desirable.

Moral encroachment exhibits a further departure from orthodoxy. Suppose we amend the Case One vignettes so the ratios are the same: The lottery has fifty tickets and one winner; there are fifty prisoners, and one refuses to riot; one in fifty iPhones thefts are committed by women, and so on. Plausibly, the evidential probability of a false verdict is the same. Orthodox epistemology holds the iPhone One verdict, and sexual assault variation, is on better epistemic footing than the Lottery One verdict. Gender may have caused the theft, for instance, or Jake's committing the sex crime is more normal. Error is more modally distant.[35] There are more background facts to draw on in forming the belief, and social features underwrite modal and causal properties. These features, which are lacking in Lottery One, can contribute to a verdict's positive epistemic standing. But moral encroachment holds that the social context – the very features that generate those positive epistemic properties – are also the features that raise the stakes or render error possibilities salient. According to some varieties of moral encroachment, the absence of rich social features is why the Lottery One verdict is on better epistemic footing – that is, more justified – than the iPhone One and sexual assault verdicts. Moral encroachment thereby reverses their respective epistemic justification ranking, as compared to orthodox views.

7. Conclusion

I examined epistemic contours of the lottery, prisoner, taxi, and iPhone vignettes, and the sexual assault variation of the iPhone vignette. (And, recall, I remain neutral about whether the judgement described in each vignette is justified.) Differences amongst the examples emerge. Some of these differences challenge Pritchard's claim that the same epistemic apparatus – his safety condition – diagnoses

[35] Thomson op. cit. note 2; Smith op. cit. note 2; Pritchard op. cit. note 4; Martin Smith *Between Probability and Certainty: What Justifies Belief* (Oxford: Oxford University Press, 2016).

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

the epistemic problem with the outright judgements based on bare statistical evidence in Case One of each pair. Pritchard argues the judgement in Lottery One, Prisoner One, Taxi One, iPhone One, and the sexual assault example are all unsafe. I have cast doubt on this claim. I argued that Pritchard's safety account does not impugn the verdict in the iPhone and sexual assault cases, and perhaps even in the prisoner and taxi examples. His account thus lacks resources to resolve the profiling paradox and proof paradox, even if it can explain the lottery paradox.

Perhaps more centrally I have shown that although the pairs of vignettes are superficially similar, they are nonetheless significantly different from each other. Even the taxi and prisoner examples used to motivate the proof paradox itself, which are often implicitly or explicitly treated as epistemically isomorphic, exhibit significant epistemic differences. The various paradoxes, when compared, show that verdicts based on bare statistical evidence are not a unified kind. These differences threaten some attempts at a unified account, such as Smith's and Pritchard's. In section six I raised and explained the question of whether theorists should aim for a unified account of the vignettes, or whether they are epistemologically too heterogenous.

I end by articulating some caution about the casual use of vignettes, especially those that stipulate statistical facts. There can be a lot going on beneath the surface of a vignette, which can threaten the probative value of the example. They can seem like clear comprehensible examples, when they are not. I think this is especially true when we stipulate, speculate, estimate, or otherwise articulate statistics or beliefs about statistics. Most features of cases, and almost all evidence, is not bare statistical evidence. If we casually stipulate statistics, we overlook and obscure important epistemic features of examples.[36]

[36] I am grateful to Mike Ashfield, Fabrizio Cariani, Jon Garthoff, Brian Hedden, Duncan Pritchard, Sherri Roush, Martin Smith, and Timothy Williamson for comments on earlier drafts. Thanks also to Kristoffer Ahlstrom-Vij, Jelena Aleksic, Boris Babic, Sven Bernecker, Renee Bolinger, Lara Buchak, Ilya Chevyrev, Julien Dutant, Amy Flowerree, Alan Hájek, Mareile Håland, Michael Hannon, Seth Lazar, Chad Lee-Stronach, Clayton Littlejohn, Silvia Milano, Zoltán Molnár, Sarah Moss, Andy Mueller, Julia Staffel, Katie Steele, and Brian Talbot for helpful discussions on these topics. I am grateful to audiences at the Ethics and Belief workshop at the University of Cologne, the Ethics and Risk workshop at the Australian National University, the Institute of Philosophy at the University of London, University of Edinburgh, and Oxford University.

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press

GEORGI GARDINER (gardiner@utk.edu) *is an Assistant Professor of Philosophy at the University of Tennessee. Before that she was the Andrew Fraser Junior Research Fellow at St. John's College, Oxford University. Her research focuses on epistemology, social philosophy, and meta-philosophy.*

Finally, many thanks to Sherri Roush's epistemology seminar at UCLA for fruitful discussions.

https://doi.org/10.1017/S0031819120000030 Published online by Cambridge University Press