Statistical evidence and the criminal verdict asymmetry

Avital Fried[1]

Abstract
Epistemologists have posed the following puzzle, known as the proof paradox: Why is it intuitively problematic for juries to convict on the basis of statistical evidence and yet intuitively unproblematic for juries to convict on the basis of far less reliable, non-statistical evidence? To answer this question, theorists have explained the exclusion of statistical evidence by arguing that legal proof requires certain epistemic features. In this paper, I make two contributions to the debate. First, I establish the Criminal Verdict Asymmetry, a previously-unarticulated asymmetry between the epistemic norms of guilty and not guilty verdicts. I argue that the prosecution and defense's different epistemic burdens influence whether statistical evidence can generate the type of verdict each side pursues. Second, I point out a mistake in how theorists have understood the role of statistical evidence in criminal trials. Though epistemologists have primarily focused on whether statistical evidence can generate specific epistemic features required for convictions, I consider whether statistical evidence can demonstrate a lack of such features. I find that there are epistemic advantages to allowing the defense to introduce statistical evidence which can reveal the prosecution's failure to prove the defendant's guilt.

1 Introduction

Otis Boone, a young Black man, was accused of committing two robberies in the United States.[1] Both robbery victims were White. In the police line-up, the police brought in Boone along with a few other people who matched the physical description provided by the victims. Both victims identified Boone as the person who had stolen

---

[1] All future mentions of this case are also from Southall (2019).

✉ Avital Fried
avital.fried@trinity.ox.ac.uk

[1] Philosophy Faculty, University of Oxford, Oxford, UK

their phones, and there was no physical evidence in the case. He was convicted and sentenced to 25 years in prison. On appeal, Boone was granted a retrial in which he successfully demonstrated to the jury that he had been many miles away from both crime locations at the times when each respective robbery took place. After serving seven years in prison, he was finally exonerated. (Southall, 2019).

Eyewitness misidentification landed Boone in prison for crimes he did not commit. The mistakes in Boone's case are potential examples of the 'cross-race effect'. The 'cross-race effect' is the psychological phenomenon that people often misidentify people of other races (Connelly, 2015). In fact, approximately 70% of those exonerated through DNA evidence in the United States were originally convicted due to witnesses who identified the wrong assailant (Agathocleus, 2020).

But whether or not juries may be warned about the 'cross-race effect' when evaluating evidence during trial is debated (Connelly, 2015; Johnson, 1983; Loftus, 1980; Wilson et al., 2013; Woocher, 1976).[2] Indeed, in Boone's case, the court prohibited the jury from hearing about the 'cross-race effect' since it constituted introducing 'bare statistical evidence', also known as naked statistical evidence.[3]

Though a few scholars support the inclusion of statistical evidence across the board, the more commonly-held view is that there is something troubling about statistical evidence in legal contexts, and especially so in the context of criminal trials. (Allen, 1990; Blome-Tillmann, 2015, 2017, 2020; Bolinger, 2020; Cheng, 2012; Cohen, 1977; Colyvan et al., 2001; Dahlman, 2020; Dant, 1988; Di Bello, 2019, 2020; Enoch & Fisher, 2015; Enoch et al., 2012; Fienberg & Schervish, 1986; Gardiner, 2018, 2020; Koehler, 2002; Moss, 2018, 2021; Nance, 2016; Nesson, 1984; Nunn, 2015; Pardo, 2005, 2010, 2018, 2019; Pundik, 2008, 2011, 2021; Redmayne, 2008; Smith, 2018; Thomson, 1986; Tribe, 1971; Twining, 1980; Wasserman, 1991; Wright, 1988;).[4] The concerns regarding statistical evidence in legal contexts fall roughly into two categories: epistemic and moral. There is a further split between the types of arguments presented regarding statistical evidence: philosophers tend to discuss the sufficiency of statistical evidence for conviction, while legal scholars tend to discuss the admissibility of statistical evidence in court.[5] This paper primarily considers *epistemic* arguments regarding the *sufficiency* of statistical evidence in criminal trials.

In this paper, I make two connected contributions to the debate on the exclusion of statistical evidence in criminal cases. First, I establish a previously unarticulated asymmetry between the epistemic norms of guilty and not guilty verdicts. This asymmetry, which I term the **Criminal Verdict Asymmetry**, reveals that the defense and prosecution have different aims—while the prosecution aims to demonstrate that the

---

[2] See, for example, Fradella (2007) for an overview of the inadmissibility of such evidence in different legal jurisdictions in the United States. See, for example, Abshire and Bornstein (2003), Johnson (1983) and Wilson et al. (2013) for more on cross-race identification and evidence admissibility in courts.

[3] The precise definition statistical evidence is debated. Some scholars take a 'I'll know it when I see it' approach. See, for example, Enoch and Spectre (2019, p. 184). Roush (forthcoming) provides the following straightforward definition: "Naked statistic is a frequency probability for a population, a mere correlation between two properties."

[4] This is by no means an exhaustive list of all the literature about statistical evidence in legal contexts. For an overview, see: Ross (2020).

[5] For a more robust discussion of these two types of distinctions, see: Dahlman and Pundik (2021).

defendant is guilty, the defense only aims to demonstrate that the prosecution has failed to prove that the defendant is guilty. The defense and prosecution, therefore, have different epistemic burdens, which could influence whether statistical evidence can generate the type of verdict each side pursues.

Second, I argue that statistical evidence introduced by the *defense* can serve an important role in criminal trials. Epistemologists have primarily focused on whether statistical evidence can generate specific epistemic features required for convictions. There is wide-spread consensus that statistical evidence alone is insufficient for legal proof. Theorists, however, have not yet considered whether statistical evidence can helpfully demonstrate a *lack* of legal proof. This type of statistical evidence, given its role, can be understood as 'debunking statistical evidence'.[6] Drawing on the **Criminal Verdict Asymmetry**, I argue that defense-introduced debunking statistical evidence can be sufficient for an acquittal.

In Sect. 2, I lay out the legal epistemological puzzle of statistical evidence, known as the proof paradox. In Sect. 3, I develop and explain the **Criminal Verdict Asymmetry**. In Sect. 4, I present a Gettier case to demonstrate how statistical evidence can reveal that judgments regarding certain propositions are unreliable. In Sect. 5, I explain how statistical evidence in criminal cases could assist the defense in showing that the prosecution has failed to prove the defendant's guilt. In Sect. 6, I raise and address two concerns.

2 Proof paradox

First, let us establish the puzzle of statistical evidence in legal contexts, known as the proof paradox. Consider the following case, based on an example first introduced by Charles Nesson (1979):

> **Prison Yard**: 100 people are exercising in a prison yard, when 99 of them suddenly join together in a planned attack on the prison guards. The remaining person tries to stop the attack. There is no available evidence distinguishing the innocent person from the rest. Local prosecutors randomly select one of them, Polly, and bring her to trial for participating in the attack.

In this case, there is no way to distinguish between those who rioted and the one person who did not participate. There is a 99% chance that Polly participated in the riot. Many people, however, share the strong intuition that it would be problematic to convict Polly solely based on that statistical evidence. Indeed, the rules of evidence in many jurisdictions would prohibit using statistical evidence in this way.[7] The exclusion of statistical evidence in criminal cases, however, is puzzling since legal systems often permit other far less reliable forms of evidence, like eyewitness accounts.

Why is it intuitively problematic for juries to convict on the basis of statistical evidence? To answer this question, theorists have explained the exclusion of statistical evidence by arguing that legal proof requires certain epistemic features, such as

---

[6] I thank one of the anonymous reviewers at *Synthese* for suggesting the term 'debunking'.

[7] See, for example, Fradella (2007).

sensitivity (Enoch & Fisher, 2015; Enoch et al., 2012), safety (Pardo, 2018; Pritchard, 2015), normic-support (Smith, 2018), or knowledge requirements (Moss, 2021).

A first possible solution to the paradox is that legal proof requires sensitivity, namely that if P were false, then we would not believe P (Enoch & Fisher, 2015; Enoch et al., 2012). In Polly's case, the belief that Polly is guilty does not meet the sensitivity requirement. If Polly was in fact innocent, we would still believe that Polly was guilty based on the evidence. A second possible explanation is that legal proof requires safety, namely that P is also true in nearby worlds (Pardo, 2018; Pritchard, 2015). In Polly's case, the belief that Polly is guilty is not safe since there is a nearby world—namely the world in which Polly was the one person who did not riot—in which P is false. A third explanation is that evidence needs to provide normic support for legal proof (Smith, 2018). A piece of evidence provides normic support for a proposition if a further explanation would be required if the proposition turned out to be false. Statistical evidence does not provide normic support for the belief that Polly is guilty, since if the belief were to be false and Polly was in fact innocent, no further explanation would be required.

Sarah Moss (2021) provides another possible solution. She argues that legal proof requires knowledge. If legal proof requires knowledge, the exclusion of statistical evidence can be explained since statistical evidence cannot generate knowledge.[8] To demonstrate why the statistical evidence alone cannot generate legal proof in the **Prison Yard** case, Moss points out parallels between Nesson's case and a well-known case presented by Gilbert Harman (1986):

> **Lottery**: Lottie has purchased a ticket in a small lottery. There are 100 tickets in all. The winning number has already been selected at random, but it hasn't been announced yet.

In this case, there is no way to know whether Lottie has a winning or a losing ticket. There is a 99% chance that Lottie lost the lottery. But if Lottie cannot rule out the possibility that she has the one winning ticket, Lottie cannot *know* that she lost. In both **Prison Yard** and **Lottery**, one may form a likely true belief on the basis of the available statistical evidence. Yet, as long as there is a live possibility that Polly is innocent or that Lottie's ticket is the winner, one cannot know that Polly participated in the riot or that Lottie lost the lottery.

Moss also contends that when a jury finds a person guilty, the jury asserts their findings by declaring the person guilty. The speech act is an assertion. It would be confusing for the jury to say: "The defendant might be innocent. We find him guilty." Moss argues that the issue with this assertion is that knowledge is the norm of assertion. Thus, if the jury does not know that the defendant is guilty, then the jury should not assert that the defendant is guilty.

Though this list of proof paradox solutions is not meant to be exhaustive, all of these solutions point to certain epistemic features which may be required for conviction. For

---

[8] Statistical evidence is unable to generate first-order knowledge, though it may be able to yield higher-order knowledge which I discuss below.

simplicity purposes, I will adopt the knowledge solution.[9] Along the way, however, I will point out how the other solutions similarly apply. In the coming section, I will discuss the epistemic difference between what the prosecution and defense aim to prove.

3 Criminal Verdict Asymmetry

There is an important, yet previously-unarticulated, asymmetry between the epistemic norms of guilty and not guilty verdicts, which I term the **Criminal Verdict Asymmetry**. To demonstrate the asymmetry, let us first begin with the aim of the prosecution. The prosecution aims to prove that the defendant is guilty. Since the prosecution carries the burden of proof, it needs to present enough evidence to the jury so that the jury is in a position to find the defendant guilty.

At first glance, it may appear that the defense aims to prove that the defendant is innocent. When a jury finds a person guilty, the jury declares: "We find the defendant guilty." A guilty verdict and a not guilty verdict are structurally the same. The two verdicts are parallel statements, one the affirmative and one the negative. If a guilty verdict is a declaration of guilt, similarly one may conceive of a not guilty verdict as a declaration of innocence. Indeed, if the defendant is found not guilty, the jury declares: "We find the defendant not guilty." This formulation of the verdict makes it seem that the jury is declaring that the defendant is not guilty.

I argue, however, that the defense is not trying to prove that the defendant is innocent. Instead, the defense only needs to demonstrate that the prosecution failed to prove that the defendant is guilty. The jury declaration of a not guilty verdict is best understood as follows: "(Not) We find the defendant guilty." Though the placement of the 'not' in the typical verdict declaration can be misleading, it is important to recognize that the defense does not aim to prove that the defendant is innocent. While the prosecution aims to prove guilt, the defense only aims to demonstrate that the prosecution has failed to meet its burden. Nothing additional is necessary for the defense to accomplish its goal of acquitting the defendant.

While this difference appears fairly straightforward, there are interesting implications for the epistemic burdens and the types of evidence which can yield the two types of verdicts. There is an asymmetry between the epistemic status of guilty and not guilty verdicts. On the one hand, the prosecution aims to prove the defendant's guilt. On the other hand, the defense only aims to demonstrate that the prosecution has failed to prove the defendant's guilt. This asymmetry impacts the type of evidence which can be used to produce each outcome. If the prosecution aims to prove the defendant is guilty, then statistical evidence is insufficient for this aim as demonstrated by the proof paradox. If the defense, however, is only trying to demonstrate that the prosecution failed to prove the defendant's guilt, perhaps statistical evidence could be introduced to yield a not guilty verdict. In the coming sections, I explore this possibility.

---

[9] The knowledge solution is more stringent than the other solutions since knowledge is factive. Moreover, the knowledge solution is distinct from the others since it explains what is both necessary and sufficient for verdicts, while the other views only explain what type of evidence is necessary for verdicts.

4 Statistical evidence in Gettier cases

In discussions about the use of statistical evidence in criminal cases, scholars typically focus on why statistical evidence is insufficient for criminal convictions. I argue, however, that statistical evidence can helpfully reveal the unreliability of judgments regarding a certain proposition which can be a helpful tool for the defense.

To demonstrate how statistical evidence operates relative to different propositions, let us consider the relationship between statistical evidence and potential beliefs in a non-legal case. Consider this classic Gettier-style case[10]:

> **Barnsville**: As Brenda is driving through a town, she sees a barn. As it happens, the structure she sees is indeed a barn. However, suppose that Brenda's friend, Cara, later tells her that Barnsville is known for its fake barns, since 90% of the structures which appear to be barns are in fact fake barns. The fake barns look identical to real barns from the road, but they are actually just facades.[11] Furthermore, Cara is a reliable and accurate source.

In this case, Brenda originally believed P, namely that the barn she saw was real. Cara then provided a piece of statistical evidence Q, namely that 90% of the structures which appear to be barns in that town are in fact fake barns. Once Brenda hears Q, Brenda no longer believes P and she may even believe ~ P, namely that the barn that she saw was fake.

Q is statistical evidence for ~ P. Q means that it is highly likely that the barn that Brenda saw was fake. However, one cannot know that the barn that Brenda saw was fake merely on the basis of Q. For the same reason that Lottie cannot know that she lost the lottery and that the prosecutors cannot know that Polly participated in the riot, Brenda similarly cannot know on the basis of Q that the barn she saw was fake. Indeed, in **Barnsville**, the barn that Brenda saw was actually a real barn. Thus, ~ P is not even true. But even if Brenda had seen a fake barn, one could not know ~ P on the basis of Q alone since statistical evidence by itself cannot generate knowledge.[12]

Importantly, Q provides direct evidence for a different proposition: Brenda does not know that P. Direct evidence can be understood as evidence for a proposition when the evidence entails that the proposition is true. Thus, Q is direct evidence for the proposition that Brenda does not know that P. Q indicates that Brenda's belief in P is not a safe belief since it reveals that P is not true in a close-by possible world.[13] If Brenda had looked at any given barn-looking structure in town, Brenda likely would not have seen a real barn. Thus, Q reveals that Brenda's original belief in P is not a safe belief. If Brenda's belief in P is not safe, then Brenda cannot know P if safety is a

---

[10] The fake barn case was first introduced by Ginet (1988) as an example of the type of case presented by Gettier (1963).

[11] There is nothing wrong with Brenda's perception since the real and fake barns look identical. Any person in Barnsville would have been unable to distinguish between the real barns and the fake barns.

[12] This is true of first-order knowledge, for instance knowledge of P.

[13] For the purpose of this paper, I use the term 'safety' in a broad sense, as I remain agnostic regarding the precise definition of safety which is debated extensively.

necessary condition for knowledge.[14] Q therefore entails that Brenda does not know that P. Thus, on the basis of Q, we know that Brenda does not know that P.[15]

Though Q is statistical evidence, it is crucial to consider its relation to a given proposition. In relation to the proposition that Brenda saw a fake barn, Q does not generate knowledge. However, in relation to the proposition that Brenda does not know she saw a real barn, Q serves as direct evidence since Q reveals that Brenda's belief in P is not a safe belief. We therefore know that Brenda does not know that she saw a real barn. While relative to one proposition Q cannot generate knowledge, Q does generate knowledge relative to a different proposition.

Let us quickly consider how some of the other proof paradox solutions apply in this case. As discussed above, Q reveals that Brenda's belief in P is not safe. Similarly, Q also reveals that Brenda's belief in P is not sensitive. If Brenda had seen a fake barn, she would have originally still believed P, namely that she saw a real barn. Thus, even if P was false, Brenda would still believe that P. Finally, Q does not provide normic support for P.[16] With Q in mind, one would not require any further explanation if P were false since it would be unsurprising that Brenda saw one of the fake barns. Now that we have established how statistical evidence works in a Gettier case, let us turn to the use of statistical evidence in criminal cases.

5 Statistical evidence in criminal cases

Let us return to the Boone case. To review, Boone was convicted for two robberies which he did not commit. He was identified as the assailant by both victims in a police line-up. There was no physical evidence in the case so the eyewitness identification by the victims was the only evidence which connected Boone to the robberies.

Boone and the actual assailant were of a different race than that of the witnesses who mistakenly identified Boone. Possibly because of the cross-race effect, the witnesses mistook Boone for their actual assailant. The eyewitness testimony served as evidence for the proposition that Boone was guilty of committing the two robberies. Let us clarify the relevant propositions so that we can understand the evidence for each proposition.

A: Boone is guilty.
B: The eyewitnesses correctly identified Boone as the assailant.

In this case, the prosecution aimed to prove A. B, namely the proposition that both victims correctly identified Boone as the assailant, was strong evidence for A. If the jury believed B, it is understandable that this would lead them to believe A. Furthermore,

---

[14] For the purpose of this paper, I adopt the widely held view that safety is a necessary condition for knowledge. A version of my argument, however, would similarly apply to the other views, such as those who reject luck-based knowledge. See, for example, Pritchard (2015).

[15] In this instance, statistical evidence appears to yield higher-order knowledge about what we can and cannot know in this case.

[16] Smith's notion of normic support addresses the relationship between a piece of evidence and a proposition—not the epistemic status of the proposition itself. The normic support view does not take a stance on the status of Brenda's belief that she saw a real barn. Rather, it can only be used to evaluate whether Q supports P.

since there was no other evidence in the case, B served as the primary, if not only, basis for believing A.

The defense was not permitted to introduce evidence regarding the cross-race effect because it was considered a form of statistical evidence. Though psychologists have not conclusively studied how frequently White witnesses misidentify Black suspects, let us suppose that an application of the cross-race effect could be translated into the following piece of statistical evidence:

C: 90% of White witness misidentify Black suspects.

In Boone's case, the eyewitnesses were White and Boone, the suspect, was Black. If C had been introduced, it would have provided support for ~ B, namely that the eyewitness did not correctly identify Boone as the assailant. Indeed, ~ B is true. Moreover, if the jury believed ~ B, then there would not have been evidence for A since the eyewitness identification was the only piece of evidence presented in the case. If the jury had no evidence to support A, then the jury would have likely found Boone not guilty.

C, however, is merely statistical evidence for ~ B. C means that it is likely that the White eyewitnesses misidentified Boone. However, one cannot know for certain whether Boone was misidentified on the basis of C alone. Again, statistical evidence cannot produce knowledge.[17] As we saw in the **Barnsville** case, the fact that 90% of the barns in Barnsville were fake did not mean that Brenda knew that the barn was fake, and in fact, the barn that Brenda happened to see was real. C only provides statistical evidence for ~ B, so one cannot know ~ B on the basis of C alone. On the basis of similar reasoning, courts have commonly excluded pieces of statistical evidence like C, since it cannot generate knowledge of the propositions the evidence appears to support.[18]

If the defense were to introduce C, however, it would not need to prove ~ B. It is not necessary for the defense to demonstrate that the White eyewitnesses did indeed misidentify Boone. If the defense aimed to prove that the defendant is innocent, then one may be concerned that the jury would mistakenly think they know ~ B on the basis of C. As discussed earlier, the jury cannot know ~ B on the basis of C alone. Per the **Criminal Verdict Asymmetry**, however, the defense does not need the jury to know that the defendant is innocent. The defense only needs to demonstrate that the jury does not know that the defendant is guilty. Thus, C is only necessary to demonstrate a lack of knowledge. In this case, C can demonstrate that the jury does not know B.

Indeed, C is direct evidence for the proposition that the jury does not know that B. C reveals that the jury's belief in B is not safe. In the majority of cases in which White witnesses identify Black suspects, the witnesses are mistaken. Thus, the jury cannot know that B, as the jury's belief in B is not safe. If safety is a necessary condition for knowledge, then C reveals that the jury does not know that B. Thus, C actually entails that the jury does not know that B. On the basis of C, we know that the jury does know that B.

Since B is the primary support for the jury's belief in A, namely that Boone is guilty, demonstrating that the jury does not know B influences A as well. If the jury does not

---

[17] The distinction between first-order and higher-order knowledge applies here as well.

[18] See, for example, Abshire and Bornstein (2003), Johnson (1983) and Wilson et al. (2013).

know B since their belief in B is unsafe, then the jury does not know that A since their belief in A is exclusively reliant on an unsafe belief. If the jury does not know that A, then the jury cannot and should not find Boone guilty.

Indeed, it is not the defense's goal to prove that Boone is innocent. Rather, in order to obtain a not guilty verdict, the defense only needs to prove that the jury is not in a position to know, on the basis of the evidence presented by the prosecution, that Boone is guilty. Statistical evidence, like C, can demonstrate that the jury's belief in Boone's guilt is based on unsafe evidence. If so, then the jury cannot know that the defendant is guilty and therefore should declare Boone not guilty.

The other solutions to the proof paradox generate a similar picture. As discussed, C reveals that the jury's belief in B is not safe. Similarly, C reveals that the jury's belief in B is not sensitive since if B was false, the jury would still believe that B. Indeed, in Boone's case, B was believed by the jury even though B was actually false since the witnesses misidentified Boone. Thus, the statistical evidence would have helped the jury realize that their belief in B is not sensitive. Similarly, C does not normically support B. Given C, no further explanation is required to understand why B is false. Simply put, it is unsurprising that the witnesses misidentified Boone given the cross-race effect. In contrast, consider a case in which the witnesses, the suspect, and the defendant were all of the same race. If the witnesses misidentified the suspect and the defendant was actually innocent, one would expect an explanation. In Boone's case, however, no further explanation is required since the cross-race effect explains the misidentification.

Importantly, debunking statistical evidence like C can be *sufficient* for a not guilty verdict. If the defense at Boone's trial had been permitted to introduce debunking statistical evidence like C, then the jury would have realized that their belief in Boone's guilt on the basis of B was insufficient for conviction. Even without additional evidence introduced by the defense, the debunking statistical evidence alone would have been sufficient to reach a not guilty verdict.

The **Criminal Verdict Asymmetry** reveals a difference in what the defense and the prosecution aim to prove and thus a difference in the types of evidence which can assist each side. Since the prosecution is trying to prove the defendant's guilt, the prosecution cannot use statistical evidence as demonstrated by the proof paradox. The defense, however, is merely trying to demonstrate that the prosecution failed to prove the defendant's guilt. Statistical evidence can demonstrate that certain pieces of evidence introduced by the prosecution are unreliable. If the defense can successfully demonstrate that the jury's belief is based on unreliable evidence, then the defense has accomplished its goal. Thus, while the prosecution cannot exclusively use statistical evidence to reach a conviction, the defense could exclusively use debunking statistical evidence to reach an acquittal.

6 Sufficiency and admissibility

Thus far, I have argued that there are epistemic benefits to defense-introduced statistical evidence. Though statistical evidence is not sufficient for conviction, I have argued that debunking statistical evidence could be sufficient for acquittal.

At the beginning of the paper, I put aside the questions about the admissibility of statistical evidence in criminal trials. Though this paper does not aim to mount a comprehensive argument in favor of allowing the defense to introduce debunking statistical evidence, some of the arguments addressed here do lay the groundwork for such an argument.

A primary reason why statistical evidence has been excluded from criminal trials is grounded in sufficiency concerns. As demonstrated by the **Prison Yard** case, statistical evidence on its own cannot yield a proper conviction. Even if statistical evidence is not the only piece of evidence introduced, there is another sufficiency concern: statistical evidence cannot properly take a body of evidence from one that is insufficient for conviction to a body of evidence which is sufficient for conviction.

My argument, however, reveals that statistical evidence can be sufficient for the defense's aim. As demonstrated by the analysis of Boone's case, the defense could have theoretically only introduced debunking statistical evidence in order to yield an acquittal. Moreover, statistical evidence can transform a body of evidence that appears to be sufficient for conviction into a body of evidence that is clearly not sufficient for conviction. In Boone's case, for instance, the body of evidence, namely the eyewitness identifications, appeared to be sufficient for conviction. If the statistical evidence regarding the cross-race effect had been introduced, the evidence would have revealed that the body of evidence was actually insufficient for conviction.

If the only reason to exclude statistical evidence from criminal trials is due to sufficiency concerns, then the arguments here provide strong reasons to permit defense-introduced debunking statistical evidence. Moreover, statistical evidence can helpfully reveal the unreliability of certain pieces of evidence. Together, those two components provide a pro tanto epistemic reason for permitting such evidence in court. The drawbacks of permitting defense-introduced debunking statistical evidence—including moral and practical concerns—would need to be further explored.[19] I will briefly raise and address two relevant *admissibility* questions.

6.1 Uneven acquittal rates

Could allowing the defense to introduce debunking statistical evidence lead to uneven acquittal rates based on a person's race, ethnicity, nationality, religion, or membership in another group? For instance, suppose that there is a neighborhood which is predominantly Black and thus almost every eyewitness to a crime is also Black. In local trials, when the defendant is not Black and the eyewitness is Black, the defense team introduces statistical evidence regarding the cross-race effect, thereby increasing the chance of acquittal of non-Black defendants. When the defendant is Black and the eyewitness is Black, however, the defense team does not introduce evidence of the cross-race effect since it is not relevant. As a result, non-Black defendants are acquitted at a higher rate in this neighborhood. Thus, one may be concerned that allowing debunking statistical evidence could lead to uneven acquittals rates on the basis of defendants' identities.

---

[19] See, for instance, Di Bello and O'Neil (2020) regarding a moral concern which arises when using statistical evidence.

My response is threefold. First, if the defense uses the cross-race effect as debunking statistical evidence in some cases but not others, it does not necessarily entail that there would be uneven acquittal rates. Rather, the debunking statistical evidence merely reveals that a certain piece of evidence presented by the prosecution is unreliable. The prosecution can and should introduce additional evidence if the prosecution aims to secure a conviction. Indeed, different amounts and types of evidence are required in different cases. This is true regardless of whether debunking statistical evidence may be introduced. For instance, if the prosecution has a video recording of the crime, there may be very little additional evidence required. In a different case, the prosecution may need to present multiple witnesses and forensic evidence in order to secure a conviction. Thus, the varied nature of the amount of additional evidence required is not a concern. Defense-introduced debunking statistical evidence does not automatically lead to an acquittal. Rather, it serves to check the reliability of the evidence presented.

Second, even if defense-introduced debunking statistical evidence leads to extra acquittals, those defendants would not have been properly convicted on the basis of the evidence presented by the prosecution. Indeed, in many legal jurisdictions if a person is factually guilty, yet there is insufficient evidence to convict, then the person is acquitted.

Third, it is worthwhile to note that aiming for an 'even acquittal rate' obscures a problematic underlying picture. There is no reason to prefer the current status quo. For instance, Black Americans are wrongfully convicted at a rate which is seven times higher than White Americans (Gross et al., 2017). If permitting defense-introduced statistical evidence led to Black Americans being acquitted at a higher rate than White Americans, perhaps the overall conviction rates by race would actually be more accurate. Thus, the notion of an 'even acquittal rate' is something that may not even be desirable. Rather, 'only proper convictions' should be the goal.

6.2 Prosecution-introduced debunking

Could the prosecution be permitted to introduce debunking statistical evidence to counter defense-introduce evidence? One may be concerned that the defense would introduce statistical evidence using reference classes which are most advantageous to them.[20] As a result, the statistical evidence may be true yet misleading. For instance, suppose that the eyewitness is a White woman and the defendant is a Black woman, so the defense introduces the cross-race effect. Suppose it is also the case that women correctly identify other women at a fairly high rate, despite the cross-race effect. One may be concerned that the defense would introduce the cross-race effect but not introduce statistical evidence that women properly identify other women at a fairly high rate. Perhaps, then, the prosecution should be permitted to introduce debunking statistical evidence to counter defense-introduced statistical evidence.

I have two responses. First, it is worthwhile to note that the debunking statistical evidence introduced by the defense would be governed by general rules of evidence. Thus, the defense would likely not be permitted to introduce misleading or irrelevant evidence, even if the evidence was technically true.

---

[20] I thank one of the anonymous reviewers at *Synthese* for raising this concern.

Second, it is certainly conceivable that the prosecution would be permitted to introduce debunking statistical evidence. This question, however, returns us to the debate regarding the sufficiency of statistical evidence. Since statistical evidence cannot move a body of evidence from insufficient for conviction to sufficient for conviction, prosecution-introduced debunking statistical evidence is likely less useful than defense-introduced debunking statistical evidence. That being said, the question of whether both sides or only the defense should be permitted to introduce debunking statistical evidence calls for further discussion.

7 Conclusion

In this paper, I first established a previously unarticulated asymmetry between the epistemic norms of guilty and not guilty verdicts. While the prosecution aims to prove that the defendant is guilty, the defense only aims to show that the prosecution failed to prove that the defendant is guilty. This asymmetry reveals that statistical evidence may be used differently by the opposing sides. I then argued that debunking statistical evidence can demonstrate that the jury's belief in the defendant's guilt is based on unreliable evidence. Thus, even though statistical evidence is insufficient for conviction, defense-introduced debunking statistical evidence can be sufficient for acquittal. If the legitimacy of a legal system relies on ensuring that convictions are correct, justified, not luck-based, and responsive to proper evidence, then the arguments presented in this paper provide strong epistemic reasons in favor of permitting defense-introduced statistical evidence.

Declarations

References

Abshire, J., & Bornstein, B. H. (2003). Juror sensitivity to the cross-race effect. *Law and Human Behavior, 27*(5), 471–480. https://doi.org/10.1023/A:1025481905861

Agathocleus, A. (2020). *How eyewitness misidentification can send innocent people to prison*, *innocence project*. Retrieved May 24, 2021from https://innocenceproject.org/how-eyewitness-misidentification-can-send-innocent-people-to-prison/

Allen, R. J. (1990). On the significance of batting averages and strikeout totals: A clarification of the naked statistical evidence debate, the meaning of evidence, and the requirement of proof beyond a reasonable doubt. *Tulane Law Review, 65*, 1093–1110.

Blome-Tillmann, M. (2015). 'Sensitivity, causality, and statistical evidence in courts of law. *Thought: A Journal of Philosophy, 4*(2), 102–112. https://doi.org/10.1002/tht3.163

Blome-Tillmann, M. (2017). "More Likely Than Not": Knowledge first and the role of statistical evidence in courts of law. In A. Carter, E. Gordon, & B. Jarvis (Eds.), *Knowledge first: Approaches in epistemology and mind* (pp. 278–292). Oxford University Press.

Blome-Tillmann, M. (2020). Statistical evidence, normalcy, and the gatecrasher paradox. *Mind, 129*(514), 563–578. https://doi.org/10.1093/mind/fzaa007

Bolinger, R. J. (2020). The rational impermissibility of accepting (some) racial generalizations. *Synthese, 197*(6), 2415–2431. https://doi.org/10.1007/s11229-018-1809-5

Cheng, E. K. (2012). Reconceptualizing the burden of proof. *Yale Law Journal, 122*, 1254–1279.

Cohen, L. J. (1977). *The probable and the provable*. Oxford University Press (Clarendon Library of Logic and Philosophy).

Colyvan, M., Regan, H. M., & Ferson, S. (2001). Is it a crime to belong to a reference class. *Journal of Political Philosophy, 9*(2), 168–181. https://doi.org/10.1111/1467-9760.00123

Connelly, L. (2015). Cross-racial identifications: Solutions to the "They All Look Alike" effect. *Michigan Journal of Race and Law, 21*, 125–145.

Dahlman, C. (2020). Naked statistical evidence and incentives for lawful conduct. *The International Journal of Evidence & Proof, 24*(2), 162–179. https://doi.org/10.1177/1365712720913333

Dahlman, C., & Pundik, A. (2021). The problem with naked statistical evidence. In C. Dahlman, A. Stein, & G. Tuzet (Eds.), *Philosophical foundations of evidence law.* Oxford University Press.

Dant, M. (1988). Gambling on the truth: The use of purely statistical evidence as a basis for civil liability. *Columbia Journal of Law and Social Problems, 22*(1), 31–70.

Di Bello, M. (2019). Trial by statistics: Is a high probability of guilt enough to convict? *Mind, 128*(512), 1045–1084. https://doi.org/10.1093/mind/fzy026

Di Bello, M. (2020). Proof paradoxes and normic support: Socializing or relativizing? *Mind, 129*(516), 1269–1285. https://doi.org/10.1093/mind/fzz021

Di Bello, M., & O'Neil, C. (2020). Profile evidence, fairness, and the risks of mistaken convictions. *Ethics, 130*(2), 147–178. https://doi.org/10.1086/705764

Enoch, D., & Fisher, T. (2015). Sense and sensitivity: Epistemic and instrumental approaches to statistical evidence. *Stanford Law Review, 67*, 557–611.

Enoch, D., & Spectre, L. (2019). Sensitivity, safety, and the law: A reply to pardo. *Legal Theory, 25*(3), 178–199. https://doi.org/10.1017/S1352325219000120

Enoch, D., Spectre, L., & Fisher, T. (2012). Statistical evidence, sensitivity, and the legal value of knowledge. *Philosophy & Public Affairs, 40*(3), 197–224. https://doi.org/10.1111/papa.12000

Fienberg, S. E., & Schervish, M. J. (1986). The relevance of Bayesian inference for the presentation of statistical evidence and for legal decision making. *Boston University Law Review, 66*, 771–798.

Fradella, H. F. (2007). Why judges should admit expert testimony on the unreliability of eyewitness testimony. *Federal Courts Law Review, 2*, 3–29.

Gardiner, G. (2018). Legal burdens of proof and statistical evidence. *The Routledge handbook of applied epistemology* (pp. 179–195). Routledge.

Gardiner, G. (2020). Profiling and proof: Are statistics safe? *Philosophy, 95*(2), 161–183. https://doi.org/10.1017/S0031819120000030

Gettier, E. L. (1963). Is justified true belief knowledge? *Analysis, 23*(6), 121–123. https://doi.org/10.1093/analys/23.6.121

Ginet, C. (1988). The fourth condition. In D. F. Austin & E. L. Gettier (Eds.), *Philosophical analysis: A defense by example* (pp. 105–117). Kluwer.

Gross, S. R., Possley, M., & Stephens, K. (2017). *Race and wrongful convictions in the United States*. University of California Irvine.

Johnson, S. L. (1983). Cross-racial identification errors in criminal cases. *Cornell Law Review, 69*(5), 934–987.

Koehler, J. J. (2002). when do courts think base rate statistics are relevant? *Jurimetrics, 42*(4), 373–402.

Loftus, E. F. (1980). Impact of expert psychological testimony on the unreliability of eyewitness identification. *Journal of Applied Psychology, 65*(1), 9–15. https://doi.org/10.1037/0021-9010.65.1.9

Moss, S. (2018). *Probabilistic knowledge*. Oxford University Press.

Moss, S. (2021). Knowledge and legal proof. In T. S. Gendler & J. Hawthorne (Eds.), *Oxford studies in epistemology 7*. Oxford University Press.

Nance, D. A. (2016). *The burdens of proof*. Cambridge University Press.

Nesson, C. (1984). Evidence or the event - on judicial proof and the acceptability of verdicts. *Harvard Law Review, 98*(7), 1357–1392.

Nunn, G. A. (2015). The incompatibility of due process and naked statistical evidence. *Vanderbilt Law Review, 68*, 1406–1433.

Pardo, M. (2005). The field of evidence and the field of knowledge. *Law and Philosophy, 24*, 321–392. https://doi.org/10.1007/s10982-004-4999-6

Pardo, M. (2010). The Gettier problem and legal proof. *Legal Theory, 16*(1), 37–57. https://doi.org/10.1017/S1352325210000054

Pardo, M. (2018). Safety vs. sensitivity: Possible worlds and the law of evidence. *Legal Theory, 24*(1), 50–75. https://doi.org/10.1017/S1352325218000010

Pardo, M. (2019). The paradoxes of legal proof: A critical guide. *Boston University Law Review, 99*(1), 233–290.

Pritchard, D. (2015). Anti-luck epistemology and the Gettier problem. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, 172*(1), 93–111.

Pundik, A. (2008). Statistical evidence and individual litigants: A reconsideration of Wasserman's argument from autonomy. *The International Journal of Evidence & Proof, 12*(4), 303–324.

Pundik, A. (2011). The epistemology of statistical evidence. *The International Journal of Evidence & Proof, 15*(2), 117–143. https://doi.org/10.1350/ijep.2011.15.2.373

Pundik, A. (2021). Rethinking the use of statistical evidence to prove causation in criminal cases: A tale of (im)probability and free will. *Law and Philosophy, 40*(2), 97–128. https://doi.org/10.1007/s10982-020-09389-0

Redmayne, M. (2008). Exploring the proof paradoxes. *Legal Theory, 14*(4), 281–309.

Ross, L. D. (2020). Recent work on the proof paradox. *Philosophy Compass*. https://doi.org/10.1111/phc3.12667

Roush, S. (forthcoming). Knowledge, evidence, and naked statistics. In *Externalism about knowledge*. Oxford University Press.

Smith, M. (2018). When does evidence suffice for conviction? *Mind, 127*(508), 1193–1218. https://doi.org/10.1093/mind/fzx026

Southall, A. (2019). A Black Man Spent 7 Years in Prison. Then a Court Changed the Rules on Racial Bias. *The New York Times*, 18 March. Retrieved May 24, 2021 from https://www.nytimes.com/2019/03/18/nyregion/cross-race-identification-witness.html

Thomson, J. J. (1986). Liability and individualized evidence. *Law and Contemporary Problems, 49*(3), 199–219. https://doi.org/10.2307/1191633

Tribe, L. H. (1971). Trial by mathematics: Precision and ritual in the legal process. *Harvard Law Review, 84*(6), 1329–1393. https://doi.org/10.2307/1339610

Twining, W. (1980). Debating probabilities. *The Liverpool Law Review, 2*(1), 51–64. https://doi.org/10.1007/BF03186639

Wasserman, D. T. (1991). The morality of statistical proof and the risk of mistaken liability decision and interference litigation. *Cardozo Law Review, 13*(2–3), 935–976.

Wilson, J., Hugenberg, K., & Bernstein, M. (2013). The cross-race effect and eyewitness identification: How to improve recognition and reduce decision errors in eyewitness situations. *Social Issues and Policy Review*. https://doi.org/10.1111/j.1751-2409.2012.01044.x

Woocher, F. D. (1976). Did your eyes deceive you-expert psychological testimony on the unreliability of eyewitness identification note. *Stanford Law Review, 29*(5), 969–1030.

Wright, R. (1988). Causation, responsibility, risk, probability, naked statistics, and proof: Pruning the bramble bush by clarifying the concepts. *Iowa Law Review, 73*, 1001.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.