Sensitivity, Causality, and Statistical Evidence in Courts of Law*

Michael Blome-Tillmann

Abstract

Recent attempts to resolve the Paradox of the Gatecrasher rest on a now familiar distinction between individual and bare statistical evidence. This paper investigates two such approaches, the causal approach to individual evidence and a recently influential (and award-winning) modal account that explicates individual evidence in terms of Nozick’s notion of sensitivity. This paper offers counterexamples to both approaches, explicates a problem concerning necessary truths for the sensitivity account, and argues that either view is implausibly committed to the impossibility of no-fault wrongful convictions. The paper finally concludes that the distinction between individual and bare statistical evidence cannot be maintained in terms of causation or sensitivity. We have to look elsewhere for a solution of the Paradox of the Gatecrasher.

1 The Paradox of the Gatecrasher

Civil proceedings in the UK and US legal systems are governed by the evidential standard of the Preponderance of the Evidence. According to this standard, a proposition $p$ counts as proven in court just in case the conditional probability of $p$ given the admissible and available evidence is greater than .5. More loosely, a proposition satisfies the standard just in case it is, as legal textbooks tend to put it, more likely than not or has a greater than 50% chance. For the purposes of this paper, let us formulate the standard of the Preponderance of the Evidence as follows, where ‘$e$’ denotes the admissible and available evidence in court:

\[
\text{Preponderance of the Evidence (PE)}: \quad p \text{ meets the standard of proof in civil suits iff } P(p|e) > .5.
\]

On this definition, a proposition $p$ has been proven to the standard of the preponderance of the evidence just in case the probability of $p$ given the admissible evidence $e$ is greater than .5. Thus, in order to find a defendant liable, the

*I am greatly indebted to Philip Ebert, Peter Graham, Alicia Hinarejos, Iwao Hirose, Jesper Kallestrup, Aidan McGlynn, Erich Schwitzgebel, Levi Spectre, Sarah Stroud, and Han-Ru Zhou for comments on earlier versions of this paper. Special thanks are due to Duncan Pritchard, Martin Smith, and two anonymous referees for this journal. Thanks also to the audiences at Edinburgh University, the Glasgow workshop on Statistical Evidence in Epistemology and the Law, and the GRIN – Université de Montréal. I am furthermore indebted to Andrew Higgins for extensive discussion and for sparking my interest in this debate. This work was supported by Marie Curie Actions (PIIF-GA-2012-328969 ‘Epistemic Vocabulary’).
The probability that the defendant is at fault must be greater than .5 given the admissible evidence. Consider an example for illustration:

*The Gatecrasher – Version A:*
The organizers of the local rodeo decide to sue John for gatecrashing their Sunday afternoon event. Their evidence is as follows: John attended the Sunday afternoon event—he was seen and photographed on the main ranks. No tickets were issued at the entrance, so John cannot be expected to prove that he bought a ticket with a ticket stub. However, a local police officer observed John climbing the fence and taking a seat. The officer is willing to testify in court.

In cases such as this one, courts will, without much hesitation, find for the plaintiff. The police officer’s eyewitness testimony establishes, in the absence of countervailing evidence, with a high-enough probability that John gatecrashed. To illustrate this further, let us assume that the police officer is, given the evidence presented in court, ordinarily reliable. Then, the probability given the evidence that John gatecrashed is .7—well above the threshold required by PE.¹

There is, however, a very similar case in which courts would find differently. Here is what I shall call *Version B* of the Gatecrasher example, which is inspired by Jonathan Cohen’s (1977: 74-81) *Paradox of the Gatecrasher:*

*The Gatecrasher – Version B:*
The organizers of the local rodeo decide to sue John for gatecrashing their Sunday afternoon event. Their evidence is as follows: John attended the Sunday afternoon event—he was seen and photographed on the main ranks during the event. No tickets were issued at the entrance, so John cannot be expected to prove that he bought a ticket with a ticket stub. However, while 1,000 people were counted in the seats, only 300 paid for admission.

In this revised version of the case, courts will find for the defendant. And again, I take it that this accords rather well with our intuitions about fairness and justice. It doesn’t seem right to find John, who was randomly picked out in the arena, liable to pay damages just because 70% of attendees gatecrashed. After all, if such a case was allowed to succeed in court, the organizers of the rodeo could, in principle, win similar cases for every person in attendance at the rodeo, including the 300 people who paid the entrance fee.

We are facing a puzzle. Why is it that courts find differently in the A- and the B-version of the Gatecrasher, given that the evidential probability that the defendant is at fault is in either case well above the threshold required by the standard of the Preponderance of the Evidence? Given that all evidence is probabilistic and that the evidential probabilities are (roughly) identical with respect to the A- and the B-version of our case, courts should find for the plaintiff in the A-version iff they do so in the B-version. But the imposition of liability on the basis of bare statistical evidence in the B-version is intuitively...

¹I assume a prior probability that the defendant is at fault of .5, which results—according to (Fields 2013: 1799)—in a posterior probability of .77. As Fields (ibid.) points out, the probability that the defendant is at fault given positive identification falls below .5 only if the prior probability that the defendant is at fault is below roughly .3. See (Schauer 2003: 317, fn. 15) for further references on the reliability of eyewitness testimony.
unwarranted and unjust. But, then, what is the decisive difference between the two different types of cases?

Cases such as the Gatecrasher and other familiar examples from the literature (e.g. the Red Cab/Blue Cab case\(^2\)) have puzzled lawyers, judges, jurists, and philosophers since at least the 1940s. A number of explanations have been proposed to resolve the puzzle. According to the most widely discussed approach to the problem, the Gatecrasher Paradox is to be resolved by distinguishing between so-called individual and bare statistical evidence, allowing for the just imposition of liability only on the basis of individual, but not on the basis of bare statistical evidence. On the approach at issue, PE is thus to be restricted to individual evidence and a resolution of the Gatecrasher is within reach: since the only evidence available in the B-version of the Gatecrasher is bare statistical or non-individual evidence, courts are not to find John liable in the B-version. In the A-version, however, the courts have convincing individual evidence at their disposal—namely, the police officer’s eyewitness testimony.

The intuitive distinction between individual and bare statistical evidence can be found in a large number of court judgments and is drawn frequently, with more or less rigour, in the legal and philosophical literature.\(^3\) But how exactly are we to understand this intuitively rather compelling distinction?

### 2 Enoch et al.’s Approach: Sensitivity

In two recent papers, David Enoch et al. (forthcoming, 2012) aim to shed light on the distinction between individual and bare statistical evidence in terms of Nozick’s (1981) notion of sensitivity. According to them, courts need sensitive evidence for the just imposition of liability. It is, as a consequence, sensitivity that distinguishes individual from merely statistical evidence, the guiding idea being that good individual evidence is sensitive, while good statistical evidence is insensitive. Sensitivity is, according to Enoch et al., the mark of individual evidence.

Before discussing Enoch et al.’s view in more detail note that, according to Nozick (1981: 179ff.), sensitivity is a property of beliefs or judgements, rather than of evidence:

\[ \textit{Sensitivity of Beliefs (SB):} \]
\[ x’s \text{ belief that } p \text{ is sensitive iff } [\text{had } p \text{ been false, then } x \text{ wouldn’t have believed } p]. \]

\(^4\),\(^5\)

However, drawing inspiration from Nozick’s account, we can formulate a similar property for evidence:

\[ \textit{Sensitivity of Evidence (SE):} \]
\[ e \text{ is sensitive to } p \text{ iff } [\text{had } p \text{ been false, then } e \text{ would have been false}]. \]

\(^2\)See, for instance, (Thomson 1986, Tribe 1971) and Smith v. Rapid Transit, Inc.

\(^3\)See (Thomson 1986: 202-203) for references. Some judges have used the terms ‘direct’ and ‘general evidence’ (see Smith v. Rapid Transit, Inc.).

\(^4\)I shall largely ignore formulations of sensitivity involving reference to belief-forming methods, but see (Nozick 1981) for discussion and fn. 6 below.

\(^5\)For more recent work on sensitivity in epistemology see (Becker and Black 2012, Cross 2010, DeRose 1995, Pritchard 2012) and (Blome-Tillmann ms).
With (SE) in place, we can finally define individual evidence along the lines suggested by Enoch et al.:

**Sensitive Individual Evidence** (SIE):

\[ e \text{ is individual evidence with respect to } p \iff e \text{ is sensitive to } p. \]

The sensitivity account of individual evidence as suggested by Enoch et al. is simple and straightforward, but it is crucially unable to solve the Gatecrasher puzzle. To see this, note firstly that, in the A-version of the Gatecrasher, the court’s evidence is in fact sensitive to the proposition that the defendant is at fault, as is illustrated by the truth of (1) in the A-version:

1. If Smith hadn’t gatecrashed, then the eyewitness wouldn’t have testified seeing Smith climb the fence.

In the A-version, the evidence—that is, the eyewitness testimony—is sensitive to the fact that John gatecrashed. To resolve the Gatecrasher puzzle, however, Enoch et al. also need to show that the evidence in the B-version is non-individual or, given (SIE), insensitive. But that is precisely not the case, as is illustrated by the truth of (2) in the B-version:

2. If Smith hadn’t gatecrashed, then it wouldn’t have been the case that while 1,000 people were counted in the seats only 300 had paid for admission.

To see that (2) is true, assume that if John hadn’t gatecrashed, then he would have paid for admission. It then follows that if John hadn’t gatecrashed, more than 300 people would have paid for admission (presumably 301). Thus, (2) is true and the court’s evidence in the B-version of the Gatecrasher qualifies—as individual evidence. Enoch et al.’s account, therefore, cannot resolve the Gatecrasher puzzle.

However, a slightly amended version of the above approach can solve this problem, and several passages in (Enoch, Spectre et al. 2012) suggest that they may in fact be attracted to this view. Consider what I shall call the belief strategy, according to which evidence is considered individual in virtue of leading to sensitive beliefs:

**Sensitive Beliefs and Individual Evidence** (SBIE):

\[ e \text{ is individual evidence with respect to } p \iff x’s \text{ belief that } p \text{ (if properly based on } e) \text{ is sensitive.} \]

Next, note that the belief strategy can resolve the Gatecrasher puzzle in a fairly straightforward way. In the A-version of the Gatecrasher, the court’s judgement (or belief) that Smith is liable is sensitive, as is illustrated by the truth of (3) in the A-version:

3. If Smith hadn’t gatecrashed, then the court wouldn’t have found him liable.

(3) is true in the A-version, because if Smith hadn’t gatecrashed, then the police officer wouldn’t have testified seeing him climb the fence and there wouldn’t
have been a case against Smith in the first place. In the B-version, however, the court’s judgement that Smith is at fault (that he gatecrashed) is insensitive, which is illustrated by the falsity of (3) in the B-version. For, if, in the B-version, the court believes that Smith gatecrashed on the basis of the bare statistical evidence available to it, then the court would still do so even if Smith hadn’t gatecrashed. In summary, (SBIE) provides us with a seemingly attractive, sensitivity-based solution to the Gatecrasher puzzle.

In what follows, I shall raise problems and produce counterexamples to this more refined sensitivity approach before turning to a discussion of causal accounts in Section 8.

3 The Opportunistic Gatecrasher

Consider the following counterexample to the above refined sensitivity approach:

The Opportunistic Gatecrasher
Sarah is on her way to her favourite pub to watch the game—as she does every Saturday afternoon. She hasn’t been able to afford the entrance to the stadium in many years, even though she’d love to watch the game there instead of in the pub. When she comes by the stadium she sees that a lot of people are gatecrashing. She decides to seize the opportunity and joins in. The evidence presented against her in court is as follows: Sarah was in the stadium (she was photographed by security cameras) and 70% of attendees in the stadium were gatecrashers. No further evidence is presented in court.

In the Opportunistic Gatecrasher, the probability that Sarah gatecrashed given the court’s evidence is again .7—just as in the B-version of the original Gatecrasher. And again, the evidence available to the court is purely statistical. However, note that this time the court’s judgement (or belief) that Sarah gatecrashed is sensitive, as is illustrated by the truth of (4):

(4) If Sarah hadn’t gatecrashed, then the court wouldn’t have found her liable.

The counterfactual in (4) is true with respect to the Opportunistic Gatecrasher because if Sarah hadn’t gatecrashed, she would have gone to the pub, wouldn’t have been photographed in the stadium, and wouldn’t have been taken to court in the first place. Thus, the court’s belief that Sarah gatecrashed is sensitive, making the evidence, according to the refined sensitivity account, individual.

Intuitively, however, courts ought not to impose liability in the Opportunistic Gatecrasher. With respect to the just imposition of liability, there is no relevant difference between the Opportunistic Gatecrasher and the B-version of the original Gatecrasher—despite the fact that the court’s evidence and beliefs are sensitive in the former, but not in the latter case. Thus, contrary to Enoch et al., sensitivity is insufficient for the just imposition of liability in courts of

---

6Of course, if Smith hadn’t gatecrashed, the statistical facts would differ slightly, for then only 699 (rather than 700) attendees would have gatecrashed. However, the court would, in the counterfactual scenario, nevertheless believe on the basis of the statistical evidence available to it that Smith gatecrashed.
law, even if construed as a constraint on the court’s judgements or beliefs rather than its evidence.7

4 The First of the Gatecrashers

Next, consider a counterexample in which the protagonist causes the evidence in question:

The First of the Gatecrashers
It’s Sunday afternoon and Hannah decides to gatecrash the local rodeo. As she climbs the fence, a large number of people in the ticket line get the same idea and follow her dubious example. More and more people start climbing the fence. Noticing that something is amiss, the organizers of the rodeo decide to take a count of the people in the stadium. Realizing that many more people are in the arena than have paid admission, they decide to take action. They randomly pick Hannah and sue her for damages. The organizer’s evidence is as follows: Hannah attended the Sunday afternoon event—she was seen and photographed on the main ranks. No tickets were issued at the entrance, so Hannah cannot be expected to own a ticket stub. However, while 1,000 people were counted in the seats, only 300 paid for admission. No further evidence is presented in court.

In this example the court’s belief that Hannah is at fault is again sensitive to the fact that she gatecrashed, as is illustrated by the truth of (5):

(5) If Hannah hadn’t gatecrashed, then the court wouldn’t have found Hannah liable.

The counterfactual in (5) is true, since in the closest world in which Hannah doesn’t gatecrash, there is no court case in the first place, because nobody at all gatecrashes, giving the organizers no reason to take Hannah to court.8

Moreover, note that any example in which the subject’s being at fault causes the statistical evidence against her will be a counterexample to sensitivity approaches to individual evidence. For if there is a causal connection between the subject’s being at fault and the existence of the evidence against her, then, if the subject wasn’t at fault, she wouldn’t have been found liable, because her

---

7One might think that this problem can be avoided by reformulating (SIBE) in terms of belief-forming methods. Assume that the belief-forming method in the Opportunistic Gatecrasher is roughly *forming a belief on the basis of photographic evidence of Sarah’s presence and the pertinent statistical evidence*. Then, the court’s belief in the example is in fact insensitive: in the closest world in which Sarah doesn’t gatecrash and in which the court forms a belief as to whether Sarah gatecrashed *on the basis of photographic evidence of Sarah’s presence and the pertinent statistical evidence* the court comes to the conclusion that Sarah did gatecrash—even though she didn’t. If, however, the belief-forming method is specified slightly differently, the problem persists. Assume, for instance, that the court forms its beliefs on the basis of photographic evidence of people in attendance and the pertinent statistical facts. Then, in the closest world in which Sarah hasn’t gatecrashed, the court won’t find her liable, because the photographic evidence is, in those worlds, of people other than Sarah. The introduction of belief-forming methods, in other words, leads to the well-known *generality problem* concerning the individuation of belief-forming methods. See (Conee and Feldman 1998) for discussion.

8For attempts to resolve this case by appeal to belief-forming methods see fn. 6.
not being at fault would have impeded the existence of the statistical evidence used against her in court. Again, the example shows that sensitivity and the individuality of evidence do not go hand in hand.

5 Necessary Truths

Beliefs in necessary truths are sensitive always and everywhere. Necessary propositions are true in all possible worlds. Consequently, for any necessary truth $p$, there is no closest possible world $w$ in which $\neg p$, and thus no possible world $w$ in which one could falsely believe that $p$. Beliefs in necessary truths are, as it were, sensitive by default. With this special property of necessary truths in mind, consider the propositions expressed by (6)-(7):

(6) John gatecrashed.

(7) John actually gatecrashed.

The proposition expressed by (6) is a contingent truth: John could have not gatecrashed. The proposition expressed by (7), however, is—given a standard interpretation of the ‘actually’-operator on which it rigidly refers to the world of the context of utterance—a necessary truth. To see this note that (7), on the intended reading, expresses the same proposition as (8):

(8) In the actual world, John gatecrashed.

Given that John gatecrashed, it follows that he gatecrashed in the actual world. The proposition that he gatecrashed in the actual world, however, is necessarily true. In any possible world, it is true that John gatecrashed in the actual world—that is, in the world that is the referent of ‘the actual world’ in our context.

But if the proposition expressed by (7) is a necessary truth, then any belief in that proposition is sensitive. Thus, in the original B-version of the Gatecrasher, the court’s judgement that John gatecrashed is insensitive, while the court’s judgement that John actually gatecrashed isn’t. Consequently, on Enoch et al.’s view, the court shouldn’t find John liable for gatecrashing, as the court’s bare statistical evidence isn’t sensitive to that charge. However, the court should find John liable for actually gatecrashing. After all, the very same statistical evidence is sensitive to that latter charge.

The problem generalizes. In any possible court case involving bare statistical evidence, courts can, according to the sensitivity approach, find the defendant liable for her actual wrongdoing, but not for her wrongdoing. As should be obvious, such a result, if unavoidable, risks reducing the sensitivity approach to absurdity, and suggests, again, that the individuality of evidence is unrelated to the modal notion of sensitivity.\(^{10}\)

---

\(^{9}\)The Red Cab/Blue Cab example comes to mind. See (Thomson 1986) for discussion and references.

\(^{10}\)As an anonymous referee points out, it might be objected here that the above problem can be avoided by demanding that individual evidence lead to non-vacuously sensitive beliefs rather than classically sensitive ones. On such an amended version of Enoch et al.’s view, however, beliefs in necessary truths can never satisfy the standard of the Preponderance of the Evidence. Such a consequence is problematic, for we can imagine cases in which experts
6 Misleading Evidence and Wrongful Convictions

A final problem to be addressed here concerns wrongful convictions. More often than desirable, courts find defendants liable that are in fact not at fault. In some such cases of wrongful conviction, the defendant is found liable on the basis of very strong but ultimately misleading evidence. In such cases, the court is not at fault. The standards of proof have been met by strong, if misleading, evidence. Call such cases no-fault wrongful convictions.

Next, note that sensitivity is factive. Necessarily, if one's belief or judgement that \( p \) is sensitive, then \( p \) is true:

\[
(FS) \quad \Box (SB[p] \rightarrow p).
\]

To illustrate this further, assume that \( x \) falsely believes that \( p \) in \( w \). Then, the closest \( \neg p \)-world to \( w \) is \( w \) itself: no world \( w \) is closer to \( w \) than \( w \). Sensitivity, however, requires that \( x \) in \( w \) not believe that \( p \) in the closest \( \neg p \)-world to \( w \), and thus, that \( x \) not believe that \( p \) in \( w \). Since \( x \) believes \( p \) in \( w \), her belief is not sensitive. Thus, sensitivity is factive: no false belief can ever be sensitive.

The factivity of sensitivity gives rise to a severe problem for the sensitivity approach to individual evidence. For if a judgement meets the standard of proof only if it is sensitive, as (SBIE) suggests, then no judgement that fails to be sensitive meets the standard of proof. But since only true judgements can be sensitive, it follows that no conviction in which the defendant is not at fault can ever meet the standard of proof of the Preponderance of Evidence. Thus, if the sensitivity approach were correct, then there couldn’t be any no-fault wrongful convictions. But there clearly are such convictions. So the sensitivity approach to individual evidence must be rejected.

7 Weakening the View?

I take it that the moral of the above problems and counterexamples is that sensitivity is irrelevant for the just imposition of liability in courts of law. The distinction between individual and bare statistical evidence cannot be elucidated in terms of sensitivity. Moreover, it is also noteworthy that, in a vast array of actual court cases, statistical evidence is used to assign liability even though that evidence is clearly insensitive. Cases of market share liability fall into this category. Counterexamples to the sensitivity approach are, therefore, not just counterfactual but can be found in actual case law.

However this may be, it might be objected at this point that Enoch et al. could, in response to the above objections, withdraw to a significantly weaker version of the views discussed so far. On such an approach, Enoch et al. are not committed to the claim that good individual evidence entails sensitivity and testify as to the truth of necessary propositions such as mathematical truths or theorems of probability theory (for instance, on the correct probabilistic interpretation of DNA evidence). Such expert testimony, however, wouldn’t meet the novel standard of evidence proposed by the referee, as our beliefs in the necessary propositions in question would be classically sensitive while failing to be non-vacuously sensitive.

\(^{11}\) Cp. (Smith 2010) for a similar point.

that good statistical evidence *entails* insensitivity—or that, necessarily, good individual evidence gives rise to sensitive judgements, while good statistical evidence doesn’t. Instead, Enoch et al. could be taken to defend the view that sensitivity and the individuality of evidence *often or usually* go together, but not necessarily. In fact, some formulations in (Enoch, Spectre et al. 2012) suggest such a softened approach:

Sensitivity-like counterfactuals capture—often enough, in sufficiently central cases—an epistemically relevant feature of the distinction between statistical and individual evidence. (Enoch, Spectre et al. 2012: 209)

While I have reservations about such a weakening of the sensitivity approach on methodological grounds, it should be noted that even the weaker view remains problematic. For if the initial Gatecrasher example is a “sufficiently central case” to give rise to a problem in our understanding of evidence that is in demand of philosophical analysis, then the same will presumably hold for the Opportunistic Gatecrasher, the First of the Gatecrashers, and the remaining cases involving ‘actually’ alluded to above. It is unclear why the normatively irrelevant details added to those novel cases should make a material difference to the just imposition of liability.\(^{13}\)

### 8 Causality and the Law

JJ Thomson (1986: 214) has proposed a view that resembles Enoch et al.’s in some respects, but which unfortunately hasn’t attracted much attention in the literature. According to Thomson, individual evidence is to be understood in causal terms:\(^{14}\)

\[(T1) \text{ } x’s \text{ } evidence \text{ } e \text{ } \text{in support of } p \text{ } \text{is } \text{individual evidence with respect to } p \text{ } \text{iff } e \text{ } \text{is caused by the fact that } p.\]

While individual evidence in support of \(p\) is caused by the fact that \(p\), bare statistical evidence in support of \(p\) is not so caused. On Thomson’s view, it is causation—not sensitivity—that is the mark of individual evidence. As Thomson (1986, pp. 208-209) puts it, it is causation that provides us with “a guarantee” that our belief that \(p\) is not merely true as a matter of luck.

While Thomson’s (T1) fares better than the sensitivity account with respect to the problem of necessary truths and remains unaffected by the Opportunistic Gatecrasher, it nevertheless remains troubled by the First of the Gatecrashers and the problem of misleading evidence and wrongful convictions. To see this, remember that in the First of the Gatecrashers the bare statistical evidence is caused by the fact that Hannah gatecrashed: Hannah caused the remaining gatecrashers to gatecrash, which brought about the statistical fact that only

\(^{13}\)An anonymous referee for this journal points out that ‘individual evidence’ is a technical term and we must therefore be offered a reliable, if not infallible guide to its meaning. Given that Enoch et al.’s account is subject to counterexamples in what are intuitively central cases, it is not clear whether they have succeeded in giving a reliable, much less infallible, guide to the meaning of the term ‘individual evidence’.

\(^{14}\)Thomson neither assumes nor defends a particular theory of causation but operates with our everyday notion of causation.
300 out of 1,000 attendees had paid their entrance fee. Thus, given (T1), the
court’s statistical evidence in the First of the Gatecrashers qualifies—contrary
to our intuitions—as individual evidence. Similarly, note that since the defend-
ant’s being at fault can cause the court’s evidence only if the defendant is in fact at fault, (T1) will not fare much better with respect to no-fault wrongful convictions than the sensitivity account did: if the defendant isn’t in fact at fault, then her being at fault cannot have caused the court’s evidence.

9 Conclusion

In light of the above problems and counterexamples, neither the sensitivity
approach nor the causal account of individual evidence seems well positioned to
resolve the Paradox of the Gatecrasher. I thus conclude that we ought to pursue
alternative approaches to that paradox, such as Smith’s (2010, ms) account of individual evidence in terms of normic support, an account suggested by Prichard (ms) that aims to understand the notion of legal risk in terms of safety rather than sensitivity, or an account that I have defended elsewhere (Blome-Tillmann forthcoming), which aims to model individual evidence within the framework of a knowledge-first epistemology.

References


