The 'She Said, He Said' Paradox and the Proof Paradox

Georgi Gardiner
University of Tennessee

This essay appears in
Truth and Trial: Dilemmas at the Intersection of Epistemology and Philosophy of Law (working title)
Edited by Zachary Hoskins and Jon Robson

Abstract
This essay introduces the 'she said, he said' paradox for Title IX investigations. 'She said, he said' cases are accusations of rape, followed by denials, with no further significant case-specific evidence available to the evaluator. In such cases, usually the accusation is true. Title IX investigations adjudicate sexual misconduct accusations in US educational institutions; I address whether they should be governed by the 'preponderance of the evidence' standard of proof or the higher 'clear and convincing evidence' standard.

Orthodoxy holds that the 'preponderance' standard is satisfied if the evidence adduced renders the litigated claim more likely than not. On this view, I argue, 'she said, he said' cases satisfy the 'preponderance' standard. But this consequence conflicts with plausible liberal and feminist claims. In this essay I contrast the 'she said, he said' paradox with legal epistemology's proof paradox. I explain how both paradoxes arise from the distinction between individualised and non-individualised evidence, and I critically evaluate responses to the 'she said, he said' paradox.

Keywords Rape accusations, Title IX investigations, proof paradox, individualised evidence, legal standards of proof, preponderance of the evidence.

1. Introduction

A 'she said, he said' case is when a third-party evaluator hears—or hears of—a rape accusation and denial, but they lack any other significant individualised evidence that bears on the case. There are no additional witnesses, compelling alibi, earlier confession, string of similar accusations, and so on. The only case-specific evidence amounts to 'one person's word against another'. But this does not mean the testimonies are epistemically balanced. In 'she said, he said' cases, probably the accusation is true. This epistemic asymmetry underwrites a paradox.

This essay describes and motivates the 'she said, he said' paradox for legal standards of proof. I first summarise some legal background. In section two I introduce the paradox. I then contrast it with the more familiar 'proof paradox' and, in section four, explain how both paradoxes arise from the distinction between individualised and non-individualised evidence. The remainder of the essay critically evaluates responses to the paradox.

Title IX investigations adjudicate accusations of sexual misconduct at US higher education institutions. The procedures determine whether university policy was violated and, if so, impose sanctions. In 2011 the Obama administration decreed that Title IX proceedings should be governed by the 'preponderance of the evidence' standard. Preponderance of the evidence, also known as the 'balance of probabilities' standard, is relatively low. The standard, which governs US civil cases involving money, is often glossed as more likely than not, or more than 50% probable, given the evidence

adduced. Betsy DeVos, education secretary in the Trump administration, countermanded the Obama-era regulation. From August 2020, institutions can choose between the preponderance standard or the more demanding 'clear and convincing evidence' standard. The clear and convincing evidence standard, which is often quantified as around 70% or 75% probability, governs civil cases involving family, such as paternity disputes and child custody, and significant life decisions such as right-to-die hearings and involuntary commitment to mental institutions. Under DeVos's guidelines, the institution must be consistent—it cannot vacillate between standards—but the lower standard is not mandated.[1]

2. Conflicting Claims

The 'she said, he said' paradox comprises six claims. Each claim is independently plausible, but they are inconsistent. At least one must be rejected.

Claim A. 'Preponderance' is the correct standard
The 'preponderance of the evidence' standard should govern Title IX proceedings for sexual misconduct hearings in US educational institutions.

Claim B. Gloss on 'preponderance' standard
The 'preponderance of the evidence' standard is satisfied if, after a well-executed investigation, given the evidence adduced, the proposition is probably true.

Claim C. 'She said, he said' evidence favours accuser
Normally in 'she said, he said' cases, based on the available evidence, the accusation is probably true.

Claim D. Considerable consequences
Finding an individual culpable of rape can warrant considerable consequences, such as expulsion.

Claim E. Liberal claim
In some cases of mere one-on-one conflicting testimony, considerable institutional consequences—such as expulsion—are not legitimatised because the evidence is mere one-on-one competing testimony from two antagonists, absent any further significant individualised evidence about the particular case.

Claim F. Connector claim
At least some conflicting testimony cases described by Claim E are normal 'she said, he said' cases, like those featured in Claim C.

The grip of the paradox is this: Suppose one hears standard 'she said, he said' conflicting testimony—a rape accusation and denial—but lacks further significant case-specific significant evidence. For at least some such cases, the following holds: Claim C says that, based on the available evidence, probably the accusation is true. So, given Claims A and B, a Title IX investigation should find the accused culpable. Given Claim D, the accused ought to face considerable consequences such as expulsion. But

[1] Simon and Mahan (1971); Gersen (2019). Gardiner (2019b) suggests the 'clear and convincing evidence' standard plausibly corresponds to knowledge-level justification.

the case qualifies as mere one-on-one conflicting testimony so—Claim E holds—for at least some such cases, considerable sanctions are not legitimated. Claim F, the connector claim, holds that such cases are possible: Claims C and E can describe the same testimonial conflict. Claim F, if true, precludes responding to the paradox by holding that 'she said' he said' cases are a different subset of testimonial conflict cases from those circumscribed by Claim E's liberal commitment.[2]

Some clarifications: Firstly, whether an accusation and denial qualify as a 'she said, he said' case can vary by evaluator. One person might only hear an accusation and denial, for example, whilst another person additionally hears a confession. Confessions are normally significant inculpatory evidence, which disqualifies the case from being 'she said, he said'. Secondly, 'she said, he said' cases are a distinctive subset of rape claims. They pick out an accused assailant, the accused learns of the accusation and responds with a denial. Accordingly they typically concern acquaintance rape, not stranger rape. Acquaintance rape is more common than stranger rape, and Title IX investigations in particular usually involve acquaintance rape accusations.[3] This essay focuses on acquaintance rape.

Thirdly, the liberal claim focuses on cases of competing testimony between involved antagonists, rather than disinterested parties, for which an institution formally sanctions one party. The liberal commitment cautions against condemnation based only on one person's word, where that person knows in advance of the accusation that no other evidence can emerge absent their change of heart and that their accusation suffices by itself for considerable consequences. It does not claim that accusations never suffice for severe sanctions; it claims they are sometimes insufficient. Fourthly, readers should think of cases, or sets of cases, that are most plausibly described by Claims A to F, rather than focusing on cases that are exceptions. As I explain below, the paradox requires only one such case. If no such cases are possible, this itself is somewhat surprising, and demands explanation.

Finally, I do not defend Claim C in this essay. Denying Claim C would resolve the paradox, but is implausible. Claim C is hard to dispute partly because it is relatively weak; it merely says such accusations are probably true.[4]

3. The Proof Paradox

The 'she said, he said' paradox may evoke legal epistemology's proof paradox. The proof paradox is generated by contrasting two vignettes, such as the following.[5]

> Gatecrasher One. A theatre sells admittance but does not issue tickets or record who paid. One day many people gatecrash. Given the number of till transactions, the manager realises 200 of the 300 attendees gatecrashed. He sues arbitrarily selected attendees for the entrance fee; Pam is one of those selected. The manager reasons that, given the base rates, Pam probably gatecrashed.

---

[2] This potential response was raised by Amia Srinivasan. I am also grateful to EJ Coffman, Jon Garthoff, and Daniel Nolan for feedback on the paradox's structure.

[3] See, for example, Lisak (2010) and Krebs et al (2007).

[4] Gardiner (ms) defends Claim C, including by arguing that false rape accusations are rare.

[5] Adapted from Cohen (1977). Other proof paradox vignettes target other legal standards, such as 'beyond reasonable doubt'. See Redmayne (2008), Gardiner (2018; 2019a), Pardo (2019) for surveys. Gardiner (2020) argues proof paradox vignettes exhibit overlooked but epistemically significant differences, and accordingly some explanations of the inadequacy of bare statistical evidence for legal proof cannot explain the full range of vignettes.

Gatecrasher Two. A theatre sells admittance but does not issue tickets or record who paid. One day a few people gatecrash. Given the number of till transactions, the manager realises 10 of the 300 attendees gatecrashed. An employee says he witnessed a particular attendee, Sam, gatecrash. The manager sues Sam for the entrance fee.

Given the available evidence, plausibly the defendant's culpability is more probable in Gatecrasher One than in Gatecrasher Two. This is because the gatecrashing base rates are far higher. The Gatecrasher Two evidence includes eyewitness testimony, but eyewitness stranger identification is notoriously unreliable.[6] Yet Gatecrasher Two evidence is the kind that can legitimate an affirmative verdict, whereas Gatecrasher One evidence cannot. This is the heart of the proof paradox.

'Quantifiable balance' conceptions of legal standards of proof claim that legal standards are numerical thresholds of probability given the available evidence. On this widely-endorsed conception, the preponderance standard is satisfied if the litigated claim exceeds 50% probability given the evidence.

The proof paradox can be constructed as four incompatible claims:

Claim 1. The 'preponderance of the evidence' standard should govern lawsuits in small claims courts.

Claim 2. The 'preponderance of the evidence' standard is satisfied if, after a well-executed investigation, given the evidence adduced, the proposition is more than 50% probable.

Claim 3. In Gatecrasher One, the evidential probability that an arbitrarily selected attendee gatecrashed exceeds 50%.

Claim 4. The evidence against Pam, an arbitrarily selected attendee in Gatecrasher One, is insufficient for finding Pam legally liable for gatecrashing.

Evidence adduced in Gatecrasher One renders Pam's guilt above 50% probable (Claim Three), yet intuitively the evidence does not suffice for an affirmative verdict (Claim Four). This challenges the quantifiable balance conception of legal proof (Claim Two).

For some theorists—myself included—'proof paradox' is a misnomer because the vignettes do not exemplify anything paradoxical; instead they simply illustrate that Claim Two is false and standards of proof are not numerical probability thresholds. As I explain in section six, these theorists can still find the 'she said, he said' paradox paradoxical because it cannot be resolved by simply denying the quantifiable balance conception of standards of proof.

4. Individualised Evidence

The proof paradox arises from the distinction between non-individualised and individualised evidence.[7] Individualised evidence, such as eyewitness testimony, is specific to features of the

[6] See, for example, Simon (2012).

[7] This distinction is controversial. Controversies include whether the distinction is tenable, how to characterise it, and what, if any, is the distinction's moral, social, or epistemic import. See, for example, Thomson (1986), Schmalbeck

particular case. Non-individualised evidence is general, and includes base rate evidence, social data, and background information about, for example, psychology and culture. Suppose S judges whether an adult, Ahmed, is taller than a teenager, Tim. Non-individualised evidence includes claims like adults are on average taller than teenagers. Individualised evidence includes evidence about their particular heights. The gatecrasher-to-payer ratio is non-individualised because it is not specific to Pam's guilt. The distinction underlies the proof paradox because, the thought goes, the court ought not find Pam culpable on non-individualised evidence alone; affirmative verdicts require individualised inculpatory evidence. One motivation for this is that non-individualised evidence is typically insensitive to the litigated claim; regardless of culpability, base rates and background social facts are about the same. Whether Pam gatecrashed makes no difference to the inculpatory evidence in Gatecrasher One. Eyewitness testimony, by contrast, is typically sensitive to culpability.[8]

The distinction between non-individualised and individualised evidence similarly underlies the 'she said, he said' paradox. Non-individualised evidence underwrites Claim C. Rape accusations generally tend to be true. Denials, plausibly, are usually false. This means the truth-to-falsity ratio in the 'rape accusation' reference class is high. Typically accusers have incentives to tell the truth, whereas deniers' incentives are frequently to lie. Deniers are more likely than accusers, as a class, to hold false beliefs about whether rape occurred. This evidence, which supports Claim C, pertains to the general class of accusations and denials. Given background base rates and social information, when a third-party evaluator hears (or hears of) an accusation and denial absent further significant individualised evidence, probably, given the available evidence, the accusation is true.[9]

Claim E is also motivated by the distinction between non-individualised and individualised evidence. In the relevant class of conflicting testimony cases, the accusation is the only significant individualised inculpatory evidence. Other than the conflicting testimonies, the individuals seem equally veracious and upstanding; there is no other individualised reason to suspect one party or the other lacks credibility.[10] Suppose the accused is expelled, for example, because rape accusations tend to be true. This means the individual is punished, in large part, based on membership of a class, such as 'those accused of rape', and the generalised features of the class—they are usually guilty. The liberal concern fears this is an illegitimate kind of profiling. Were the accused innocent, the supporting evidence would nonetheless obtain; the background evidence is insensitive to guilt in the individual case.[11]

5. Flexibility

The 'she said, he said' paradox shares various other features with the proof paradox. Both are located at the intersection of evidence and action, particularly where that action is—or is mediated by—a

(1986), Schoeman (1987), Schauer (2003), Di Bello and O'Neil (forthcoming), Bolinger (this volume), Littlejohn (this volume), and Smith (this volume).

8 Enoch, Spectre, and Fisher (2012), Blome-Tillmann (2015), and Gardiner (2018) discuss the relation between individualised evidence and sensitivity. Wright (this volume; ms) defends the view that testimony is not evidence of the truth of the asserted content.

9 Gardiner (ms) defends these claims at length.

10 Crucially the conflicting testimonies are epistemically significant; there is no *further* significant individualised evidence.

11 I grant that most accusations are sensitive to guilt: if the rape did not occur, the accusation would not have been issued. But the evidence underwriting Claim C is non-individualised and so is insensitive to guilt. The liberal concern is that if this non-individualised, insensitive evidence is why the institutional authority favours the accuser's assertion over the accused, and imposes considerable sanctions, this qualifies as profiling. In future research I will address whether the liberal concern similarly applies to other assertions, not just antagonist accusation, and whether adjacent paradoxes arise for other kinds of investigation. Thanks to Jon Garthoff and Jon Robson for helpful discussions on these topics.

formal finding of culpability. They both threaten the quantifiable balance conception of legal standards of proof. They both need only one plausible case to be efficacious. Gatecrasher One need not have happened; the thought experiment alone suggests the preponderance standard is not a quantifiable balance threshold. Similarly, the mere plausibility of a 'she said, he said' case described by Claims C and E suffices; the case need not actually obtain.

We can thus imagine cases, for example, where 'she said, he said' testimonies are mediated through testimonial transition. The Title IX investigator never interviews the accuser or accused; she simply hears second-hand reports. In this imaginary 'Transmitted Accusation' case, the accusation is probably true given the evidence available to the administrator (Claim C), and yet the administrator should not normally expel the accused based on this evidence (Claim E). Thus the paradox can arise from fictional examples. Note, though, I think the tension amongst the six claims—tension rooted in the distinction between individualised and non-individualised evidence—arises for some real-life cases.

The 'she said, he said' paradox is inherently flexible; this flexibility is a strength because Claims C and E are accordingly plausible to a wider range of readers. It stems from flexibility in what qualifies as a 'she said, he said' case, compared to an accusation and denial augmented with significant additional individualised evidence. Consider, for example, an accusation and denial plus the further fact that the accuser seemed unperturbed the day after the alleged rape. Some people judge this behaviour to be significant evidence against the accusation, and so do not regard the case a 'she said, he said' case. Others regard the behaviour as nonprobative or insignificant, and so consistent with its being a 'she said, he said' case. Regardless of who is correct about whether the behaviour is significant evidence, they disagree about which cases are described by Claim C. Here is why the flexibility is a strength: What matters is only that at least some cases described by Claim C are also cases described by Claim E. Theorists can disagree about which cases satisfy this requirement. If a theorist can think of cases, real or imaged, such that on their view both Claims C and E hold for that case, the paradox arises. This condition is articulated by Claim F, the Connector Claim.

Some theorists have a narrow conception of 'she said, he said' cases, in which few examples qualify. They hold that almost any individualised additional facts—such as character references, how the parties emote during testimony, and who can reliably recall details—count as sufficiently significant to disqualify the case from being 'she said, he said'.[12] Other theorists have a broader conception, so that more cases qualify as 'she said, he said' cases. They regard various additional individualised facts, such as emotions exhibited during testimony, as nonprobative or paltry and so consistent with the cases being 'she said, he said' cases. Other theorists might focus on one specific real-life or imaginary case, such as Transmitted Accusation. One can read Claim C as about what is typical, characteristic or normal for all 'she said, he said' cases, a subset, or at least common for a wide class of 'she said, he said' cases. There is even flexibility about whether 'she said, he said' cases are limited to bald accusations and denials or can include narrative details about what occurred.[13] What matters is only that, as articulated by the Connector Claim, in at least some such case the accusation is probably true, given non-individualised evidence, but considerable consequences are illegitimate because the inculpatory evidence is insufficient.

---

[12] As I argue in section nine and Gardiner (ms), I think the narrow conception is mistaken because most such facts are paltry or nonprobative.

[13] With thanks to Jon Robson for helpful comments about the scope of 'she said, he said' cases. Gardiner (ms) discusses the epistemological consequences of supplying further narrative details, including outlining how such details can fuel doubt.

6. Relative Strengths

In addition to the similarities, the 'she said, he said' paradox also differs from the proof paradox. I articulate two potential weaknesses of the 'she said, he said' paradox relative to the proof paradox, and then describe relative strengths. Note that describing comparative strengths and weaknesses of the two paradoxes is largely a rhetorical device to enhance understanding of the paradoxes; there is no competition between the paradoxes.[14] Firstly, insofar as the two paradoxes challenge the quantifiable balance conception of standards of proof, the 'she said, he said' paradox has a disadvantage. There are various alternative claims one might reject in order to defend the claim that the 'preponderance' standard is satisfied if the litigated claim's evidential probability exceeds 0.5. A theorist might respond by rejecting Claim A or E, for example. The proof paradox, by contrast, has fewer alternative resolutions.

Secondly proof paradox vignettes are—at least superficially—clean and straightforward. There is relatively little disagreement about how to interpret Gatecrasher vignettes. There is less agreement and clarity about precisely delineating 'she said, he said' cases. The tension underwriting the paradox can be correspondingly harder to appreciate.[15]

The proof paradox vignettes are clean, straightforward, fictional vignettes. This apparent simplicity is, in some respects, a virtue. But these features are also a weakness. They are unrealistic, artificial, pretend cases. The proof paradox is dismissed by theorists who claim such cases never actually arise and so cannot inform the epistemology of legal proof.[16] It is difficult to reconcile the lack of further evidence in proof paradox vignettes, furthermore, with the stipulation that the investigation was conducted appropriately. Realistically in Gatecrasher One, for example, security footage would be available, attendees would be questioned, and so on. The lack of further individualised evidence against Pam is suspicious and implies a flawed investigation.

In 'she said, he said' cases, by contrast, the absence of further significant individualised evidence is—in many cases—normal and expected. It is not suspicious because rape often lacks corroborating evidence; lack of other witnesses and evidence is often a pre-condition for the attack. The cases are accordingly not so easily dismissed as a challenge to the quantifiable balance conception of legal proof. That is, the 'she said he said' paradox cannot be disregarded as hinging on artificial, unrealistic vignettes or stipulations.

A second virtue of the 'she said, he said' paradox, relative to the proof paradox, is that we feel the tension for real-life cases, which fuels an urgency to resolving the paradox. It is not merely an abstract intellectual puzzle; it is a challenge society must address. Even if strict 'she said, he said' cases are rare, various responses to the paradox bear on how we should treat rape accusations that are not strictly 'she said, he said' cases. The resolution has policy implications. Some interlocutors oppose Claim D for Title IX investigations, for example, to protect Claims A and E.[17] If we jettison the liberal

---

[14] Thanks to Cécile Degiovanni for insightful comments on this topic.

[15] Gardiner (2020) argues the apparent simplicity of many proof paradox vignettes is illusive. Unclarity about how to precisely delineate the class of 'she said, he said' cases might also obscure that ultimately there is no paradox: Perhaps Claim F is false and there are no relevant possible cases. Section nine discusses this response to the paradox.

[16] See, for example, Allen (forthcoming). Enoch (forthcoming) replies. Some real-life cases are redolent of proof paradox vignettes (Redmayne (1996); Koehler (2001)), but they differ substantially, often in overlooked ways. The real-life *Smith v. Rapid Transit* evidence, for example, involves no ratios or numerical probabilities. Gardiner (forthcoming-b; 2019) contrasts *Smith v. Rapid Transit* with proof paradox vignettes.

[17] Greer (2018) argues for lowering penalties for rape to secure more rape convictions.

commitment, more accused individuals should face severe sanctions, despite the absence of significant individualised evidence beyond the accusation. Rejecting Claims A or B, however, renders rape harder to prove.

The proof paradox challenges quantifiable balance conceptions of legal standards of proof, such as that the preponderance standard is a 0.5 evidential probability threshold.[18] This is because the evidence available underwrites a high numerical evidential probability, but seems insufficient for an affirmative verdict. The 'she said, he said' paradox also threatens quantifiable balance conceptions of legal standards: In 'she said, he said' cases, the quantifiable balance surpasses 50%. But a relative virtue of the 'she said, he said' paradox is that insofar as it challenges quantifiable balance conceptions, it also challenges other relevantly similar conceptions. Claim B of the 'she said, he said' paradox holds,

> Claim B. The 'preponderance of the evidence' standard is satisfied if, after a well-executed investigation, given the evidence adduced, the proposition is probably true.

Claim B is not restricted to numerical, quantifiable balance conceptions of probability. This contrasts with the corresponding claim of the proof paradox, which is limited to a quantifiable balance conception.

> Claim 2. The 'preponderance of the evidence' standard is satisfied if, after a well-executed investigation, given the evidence adduced, the proposition is more than 50% probable.

We can resolve the proof paradox by denying the claim that the preponderance standard is satisfied if the litigated claim exceeds 50% evidential probability. We cannot, however, resolve the 'she said, he said' paradox by denying this claim; the paradox does not rely on this claim. This is because for a wide range of qualitative conceptions of evidential favouring, 'she said, he said' evidence favours the accusation.[19] It is more normal for an accusation to be true and a denial false, for example, than the converse. If an accusation is false, this is surprising; the aberration from the norm demands explanation. This suggests 'she said, he said' evidence normically supports the accusation; evidence normically supports a proposition when, roughly speaking, given that evidence, the proposition would normally be true.[20] The truth of the accusation better explains the available evidence than the truth of the denial; more plausible and simple narratives are available if the accusation is true. The accusation's truth coheres better with background facts, such as human nature, than the denial's truth. Belief in guilt from 'she said, he said' evidence is plausibly more safe than belief based on Gatecrasher One evidence; error is a more distant possibility.[21] The available evidence is also more sensitive to truth.[22]

On various conceptions of which epistemic properties justify belief, insofar as you should believe one speaker, you should believe the accuser rather than the denier. You should have higher confidence in

---

[18] Gardiner (2020) argues proof paradox vignettes threaten other theories of legal proof, such as Pritchard's (2017) safety-based account.

[19] Gardiner (ms-b) argues for the epistemic potency of rape accusations.

[20] Cf. Smith (2016; 2018), Gardiner (2018), Littlejohn (this volume).

[21] Pritchard (2005; 2017).

[22] Evidence in 'she said, he said' cases, but not in most proof paradox cases, typically tracks the truth. The inculpatory evidence in Gatecrasher One would be the same regardless of whether Pam gatecrashed. But if the accused did not rape, typically the accusation would have not have been issued. Accusations—but not their pedigree—is sensitive to truth. Cf. Roush (2005). This epistemic feature of 'she said, he said' evidence could be leveraged against the liberal commitment articulated by Claims E and F.

the accusation and, absent further evidence, be more suspicious about the denial. These properties plausibly do more than render the accusation numerically probable. They also underwrite qualitative forms of epistemic support. The 'she said, he said' paradox accordingly threatens more conceptions of standards of proof than the proof paradox can.

To summarise: We can respond to the proof paradox by rejecting Claim Two, and supplanting it with a more demanding conception of the preponderance standard.[23] Perhaps 'preponderance' requires not merely that the litigated claim exceeds 50% evidential probability, for example, but also that the evidence normically supports the conclusion. This might plausibly resolve the proof paradox. But 'she said, he said' evidence favours the accusation on various conceptions of epistemic support and so may well satisfy the revised conception of the preponderance standard. Merely denying the quantifiable balance conception of legal proof does not resolve the 'she said, he said' paradox. It is a more resilient paradox.

7. The Epistemology of Preponderance

This observation raises the question of whether there is a conception of the preponderance standard such that 'she said, he said' evidence fails to satisfy that standard. That is, can we resolve the paradox by revising Claim B, our conception of the epistemology of the preponderance standard?

One such strategy holds that satisfying a legal standard of proof requires producing independent corroborative inculpatory evidence. Another strategy avers that legal proof requires ruling out relevant alternatives, and—crucially for this strategy—'she said, he said' evidence leaves relevant alternatives uneliminated.[24] The rape is not proven to a preponderance of the evidence, on this view, unless the evidence addresses all relevant error possibilities, and some relevant error possibilities are not addressed by rape accusations alone. One way to develop this view holds that if two involved antagonists make conflicting claims, it is a relevant alternative that one is simply lying when they report to the third party. On this conception, 'she said, he said' evidence fails to satisfy the preponderance standard because it does not address this error possibility. Further evidence, such as the accuser's contacting her therapist immediately afterwards, would address this possibility. Given this further evidence, it is still possible that she is lying, but only if she also deceives the therapist; the 'simple lie' error possibility is addressed because the only remaining uneliminated error possibilities involve addition deceptions.[25]

This view might support the claim that 'she said, he said' evidence fails to satisfy the preponderance standard, but as stated the conception of the preponderance standard is implausibly demanding. If one antagonist claims something impossible or wholly implausible, this does not thereby render relevant the possibility that the other is lying. Suppose the defendant in a vandalism lawsuit asserts the plaintiff's car was smashed by aliens, and the plaintiff denies this. The mere fact of conflicting testimony between antagonists does not render relevant that the plaintiff is lying about the absence of alien involvement. The defendant's claim is too farfetched to take seriously. The strategy, then, must

---

[23] This strategy is pursued in, for example, Thomson (1986), Ho (2008), Enoch, Spectre, and Fisher (2012), Pritchard (2017), Moss (2018; forthcoming), Smith (2018), Gardiner (2019b), and Littlejohn (this volume). For surveys, see Redmayne (2008), Gardiner (2019a), and Pardo (2019).

[24] Ho (2008), Amaya (2015), Moss (2018; forthcoming), and Gardiner (2019; forthcoming-b) develop relevant alternatives conditions for legal standards of proof.

[25] For discussion, see Gardiner (2019b, forthcoming-a).

be refined. Not all antagonist disagreement can by itself render lying relevant. The refinement must walk a tightrope: It must articulate constraints on when conflicting antagonist testimony itself renders lying relevant; but they cannot be too restrictive if the accuser's lying is a relevant alternative in all or typical 'she said, he said' cases.

I endorse neither the corroboration requirement nor the relevant alternative strategies for resolving the 'she said, he said' paradox by denying Claim B.[26] Note such strategies render rape harder to prove because, as noted above, there is often little substantial corroborative evidence. Whether this is an objection depends on the aims of legal institutions. If rape is difficult to prove, is this a problem with the epistemological features of legal proof, a fault within social or legal institutions, or simply one of the tragic features of the crime?

Thus it appears challenging to resolve the 'she said, he said' paradox by denying that 'she said, he said' evidence satisfies the preponderance standard. This is because on various conceptions of epistemic support that might characterise the preponderance standard, 'she said, he said' evidence can satisfy it. Indeed, one way to understand the crux of the 'she said, he said' paradox is the contrast between how epistemically forceful rape accusations are compared to the relatively low preponderance standard; accusations are so epistemically potent that they, by themselves, seem to satisfy the standard, at least on many construals of 'preponderance'. The paradox arises because this threatens our liberal commitment against condemning individuals based on one person's say so, at least in some cases in which the accuser is an involved antagonist rather than a disinterested party, and where she antecedently knows no independent exculpatory evidence can come to light absent her change of heart, and where she knows her accusation alone suffices for severe sanctions.[27]

8. Higher Standards

Perhaps rejecting Claim A also cannot resolve the paradox. Gardiner (ms) epistemically evaluates acquaintance rape accusations. I focus on three aspects: Track records of truth and falsity, incentives to lie, and which speaker typically possesses better evidence.[28] This evaluation supports Claim C—in 'she said, he said' cases, accusations are probably true. Claim C is relatively weak; it merely concerns what is more probable. But plausibly this non-individualised evidence renders accusations considerably more probable than denials. And so one might wonder whether accusations are so strongly indicative of rape—accusations are so likely true—that 'she said, he said' evidence also satisfies higher standards of proof, such as the 'clear and convincing evidence' standard.

Suppose one holds that Title IX investigations should be governed by the clear and convincing evidence standard. Ostensibly this view escapes the paradox—it denies Claim A. But acquaintance rape accusations are rarely false. Kelly et al (2005) suggests false rape claims tend to exhibit distinctive features—they describe violent stranger rape and do not accuse specific individuals, for example, or accuse government officials for financial motives. As Newman (2017) describes, false rape claims tend to look very different from acquaintance rape accusations. Plausibly, by contrast, acquaintance rape

[26] Gardiner (2019b) develops a relevant alternatives account of legal proof, but does not discuss whether 'she said, he said' evidence characteristically leaves relevant alternatives uneliminated. Gardiner (forthcoming-a) applies the relevant alternatives account to the epistemology of rape accusations and investigates when and why the accuser's speaking falsely is a relevant alternative.

[27] Gardiner (ms: §8) defends the liberal commitment; this two-pronged defence highlights epistemic weaknesses of accusations and socio-political motivations.

[28] These aspects are not independent; incentives to lie underwrite high false denial rates, for example.

denials are typically false.[29] If so, perhaps 'she said, he said' evidence also satisfies the clear and convincing evidence standard. Correspondingly, an adjacent paradox arises:

Claim A'. 'Clear and convincing evidence' is the correct standard
The 'clear and convincing evidence' standard should govern Title IX proceedings.

Claim B'. Gloss on the 'clear and convincing evidence' standard
The 'clear and convincing evidence' standard is satisfied if, after a well-executed investigation, the total evidence adduced clearly and convincingly supports the proposition.

Claim C'. 'She said, he said' evidence is clear and convincing
In 'she said, he said' cases, the total available evidence (commonly, characteristically, normally, or at least sometimes) clearly and convincingly supports the proposition.

Claim D. Considerable consequences
Finding an individual culpable of rape can warrant considerable consequences, such as expulsion.

Claim E. Liberal claim
In some cases of mere one-on-one conflicting testimony, considerable institutional consequences—such as expulsion—are not legitimatised because the evidence is mere one-on-one competing testimony from two antagonists, absent any further significant individualised evidence.

Claim F'. Connector claim
At least some conflicting testimony cases described by Claim E are normal 'she said, he said' cases, like those featured in Claim C'.

Some readers endorse Claim C'. They aver that 'she said, he said' cases normally, or at least sometimes, provide clear and convincing evidence. Arguably the epistemic force of the maxim to believe accusers, recently articulated by the hashtag #BelieveWomen, is that rape accusations are—by themselves and despite denials—clear and convincing evidence. On this view, they are believable because an impartial, objective epistemic evaluation of the epistemology of rape accusations finds them clear and convincing.[30] But Claim C' generates a 'she said, he said' paradox for the higher DeVos standard.

Recall the quantifiable balance conception of the clear and convincing evidence standard. On this view, if the litigated claim is at least, say, 70% or 75% probable given the evidence, the standard is satisfied. This means that—if rape accusations have a sufficiently high truth-to-falsity ratio, and corresponding denials are often enough false—'she said, he said' cases typically satisfy the standard.[31]

This reasoning suggests a further argument against the quantifiable balance conception of legal standards. Suppose you think beyond reasonable doubt is a quantifiable balance threshold, such as

[29] Gardiner (ms) defends these claims. Cf. Jordan (2001), Kelly et al. (2005), and Lisak et al (2010). Note that stranger rapes are often not, strictly speaking, accusations: No individual is accused.

[30] Ferzan (2021) and Bolinger (2021) discuss the epistemology of #BelieveWomen.

[31] It is possible that accusations are almost always true, yet denials are also always true. This is because not all accusations lead to denials. Perhaps accusations that lead to denials are disproportionately false.

90% or 95% evidential probability. If the truth-to-falsity ratio amongst accusations is sufficiently high, and amongst denials sufficiently low, then normally 'she said, he said' evidence renders the accusation above 90% or 95% probable. And so, given the quantifiable balance conception, 'she said, he said' evidence, including in cases like Transmitted Accusation, satisfies the reasonable doubt standard. But this seems implausible. Even if 'she said, he said' evidence sometimes proves guilt beyond reasonable doubt, Transmitted Accusation evidence cannot: Criminal culpability cannot be established merely by membership in a sufficiently inculpatory reference class.

I think rape accusations usually are true, warrant belief, and are compelling evidence; Gardiner (ms) argues at length for their excellent epistemic standing. But it seems dogmatic or unimaginative to say 'she said, he said' evidence can, since non-individualised base rate evidence is sufficiently supportive, provide proof beyond reasonable doubt.[32] The reference class—accusations—strongly indicates truth, but there is nonetheless sometimes room for reasonable doubts. As with proof paradox vignettes, comparing 'she said, he said' cases to the reasonable doubt standard suggests reference class membership cannot by itself satisfy beyond reasonable doubt, and accordingly the legal standard is not a quantifiable balance threshold.

9. Doubting the Category

I conclude by evaluating two other responses to the original 'she said, he said' paradox. The first denies the 'she said, he said' category is cogent. 'She said, he said' cases, recall, are accusations of rape, followed by denials, with no further significant individualised evidence available to the evaluator. But, as I noted in section five, further evidence influences evaluation of testimony. We have antecedent background beliefs about how plausible or likely the content is, which topics and contexts beget honest assertions and accurate beliefs, and about speaker veracity, either for individuals, members of reference classes, or testimony in general. We harbour background beliefs—often unacknowledged, overconfident, or inaccurate—about the conduct of 'real' rape victims and perpetrators.

These background beliefs influence how third-parties evaluate rape accusations and denials, and underlie disagreement about which cases qualify as mere 'she said, he said' cases, as opposed to cases with further significant individualised evidence. Evidence deemed significant by some is deemed paltry or nonprobative by others. This is commonly seen with, for example, accusations when the accuser was friendly to the accused after the incident in question. Many third-party evaluators consider this not a 'she said, he said' case because they regard the friendliness as significant exculpatory individualised evidence. Other third-party evaluators regard this conduct as insignificant evidence—or not evidence at all—against the accusation because it is normal for acquaintance rape victims to be friendly to their assailant afterwards. They can regard it a 'she said, he said' case. Such disagreements might suggest we cannot make sense of 'she said, he said' cases as a class, because the definition includes the lack of 'further significant individualised evidence', but background beliefs play an ineliminable role in evaluating testimony; we cannot isolate the epistemic role of testimony from that of background beliefs.

I do not find this response to the paradox compelling. Firstly, epistemological vignettes often stipulate that the central or only evidence is testimonial; the epistemic role of background beliefs in such

[32] I think uncorroborated rape accusations warrant belief, but the 'she said, he said' paradox needs only the weaker Claim C. Note also it would be welcome news, for feminist reasons, if my reservations are unduly cautious and rape accusations are by themselves epistemically potent enough to satisfy beyond reasonable doubt.

vignettes does not typically impede theorising about the epistemic force of testimony. The relative chariness about the 'she said, he said' category may emerge from an encultured underestimation of the epistemic force of rape accusations relative to the perceived importance of background facts such as conduct and character. This overestimation of the relative significance of nontestimonial evidence when evaluating rape accusations leads us to think we cannot cleave the epistemic force of such accusations from an evaluation of whether, for example, the accuser seems sufficiently upset.

Secondly, most such supplementary facts, such as the accuser's friendliness or asserters' mannerisms, are consistent with 'she said, he said' cases; they are nonprobative or not highly probative, but evaluators erroneously regard them as significant evidence. Given this, we need not hand-wring over how precisely to delineate 'she said, he said' cases. Claim C is likely true even under a broad conception of the category. Indeed, we are likely unreliable or anti-reliable at estimating the probative force of such evidence. Narrative gaps and inconsistencies, fidgeting, evasiveness, irresolute emotions and reactive attitudes, not realising or reporting until much later, and so on are taken to be individualised evidence of false accusations, for example, but are symptoms of trauma.[33] On hearing the paradox, some people emphasise that real-life cases exhibit additional evidence, such as mannerisms; the implication is that real-life cases are never bare 'she said, he said' cases. I think this response overestimates our reliability at discerning truth in such contexts.

Thirdly, as argued above, the paradox permits congenial flexibility about which cases qualify; what matters is consistency within the paradox, such that at least some cases—or class of cases—are described by both Claims C and E.[34] Fourthly, one can recreate the paradox by focusing on a single accusation and denial, such that Claims C and E describe the case. The case could be imaginary, like Transmitted Accusation, or a concrete real-life case. Focusing on a single case circumvents conceiving of cases via the 'she said, he said' category. So long as the claims hold for at least one example—actual or merely possible—the paradox arises. And I think such cases will arise: As noted earlier, I think the tension between the claims is felt for some real-life cases. Finally, Claims C and E stem not only from intuition about real-life cases but also from principled reflection on the epistemic and social importance of the distinction between individualised and non-individualised evidence. This distinction itself generates the possibility of accusations that are epistemically well-supported by non-individualised evidence, but that do not legitimate considerable consequences because the only individualised inculpatory evidence is isolated antagonist testimony. For these reasons, I doubt demurral about precisely delineating the 'she said, he said' category can evade the paradox.

10. Title IX Overhaul

Various people suggest the following response to the paradox: US educational institutions should abandon the existing Title IX system, which is loosely based on legal adjudication. The Obama standard and DeVos revision apply to a system in which complainants, their representatives, or an institutional authority must establish a claim to a particular epistemic threshold before imposing sanctions. Interlocuters suggest that for sexual misconduct accusations in US educational institutions,

[33] Thanks to Sarah Robins and Hilary Kornblith for discussion.

[34] Philosophy generally requires charitable interpretation of cases. As Williamson (2007) notes, we can imagine ways of filling out Gettier cases, for example, which fail to constitute a counterexample to the target theory. But theorists instead interpret the vignettes nondeviantly. Responding to Williamson, Gardiner (2015) emphasises the ineliminable role of normal interpretations of vignettes for philosophical theorising. Similarly we should focus on 'she said, he said' and proof paradox examples that can—rather than fail to—illustrate the tension among Claims A–F and 1–4.

this format should be replaced with a model that does not require standards of proof.[35] One proposed rival model is restorative justice, which focuses on rehabilitation, reconciliation, and repairing harms, rather than punishing or removing the offender from the relevant community. Others propose a private industry model, in which authorities can typically respond as they see fit—demotions, moving departments, adjusting schedules or responsibilities, termination of employment, and so on—without first needing to prove claims to a particular epistemic threshold. These responses to the paradox deny Claim A; they aver Title IX investigations should not be governed by the preponderance standard. But rather than propose a higher standard, they overhaul the system entirely.

In response, these rival approaches themselves face problems. Restorative justice might be unsatisfying to victims, for example, or be insufficiently consequential given that other campus misconduct, such as non-sexual assault, elicits orthodox sanctions. Educational institutions are a public good and play distinctive structural roles in society, including bestowing qualifications and imprimatur. These features might make the private industry model inappropriate for sanctions such as expulsion. That is, plausibly a student has a greater right to remain enrolled in college, despite accusations, than to remain in private employment in light of accusations; more formal procedures are needed to expel a student.

The paradox is not easily evaded. Suppose Title IX procedures at US educational institutions are revised to eschew standards of proof. This might be a practical response to the paradox, but the tensions underlying the paradox remain. The tensions arise from general features of testimony, legal proof, epistemic thresholds, sanctions, rights, and the distinction between individualised and non-individualised evidence. Adjacent paradoxes can thus arise for various legal and institutional contexts, such as child custody, civil damages, revoking licences, employment tribunals, and criminal conviction. Variations of the paradox may even arise for interpersonal relationships, such as friendships. I conclude by highlighting three questions at the core of the 'she said, he said' paradox.

Firstly, how can we reconcile the potent epistemic force of rape accusations—they are excellent evidence of guilt—with our liberal commitment against condemnation by one person's say so when the accuser knows in advance that no other evidence can come to light absent her change of heart and that her accusation alone suffices for serious sanctions?

Secondly, how can we reconcile a conception of standards of proof as quantifiable balance thresholds—or even thresholds of qualitative epistemic properties such as safety—with our judgement that some kinds of evidence cannot by themselves satisfy those standards? After all, non-individualised evidence, such as base rates and background knowledge about psychology and culture, render the litigated claims highly probable, safe, and so on. So why doesn't this evidence alone legitimate affirmative verdicts?

Finally, how can we reconcile the feminist tenet that findings of culpability for rape warrant considerable consequences, and should express weighty social meaning, with the claim that such findings should be governed by the relatively low preponderance standard? One might object that expulsion and termination of employment are not severe sanctions. I grant they are not severe relative to the crime of rape. But they are relatively severe when contrasted with the tenuity of the preponderance standard. Preponderance, bear in mind, is usually glossed as merely more likely than not.

[35] Thanks to Clayton Littlejohn and Alannah Tomich for helpful discussions.

These unresolved tensions form the heart of the 'she said, he said' paradox.

Acknowledgements
Many thanks to Richard Arning, Anne Baril, Renee Bolinger, Becky Brown, Bruce Chapman, Ilya Chevyrev, Cécile Degiovanni, Julien Dutant, Catherine Elgin, Bryan Frances, Jon Garthoff, Alan Grafen, Michael Hannon, John Hardwig, David Henderson, Zachary Hoskins, Torfinn Huvenes, Kareem Khalifa, Hilary Kornblith, Hafsteinn Kristjánsson, Barry Lam, Clayton Littlejohn, Jonathan Livengood, Tim Loughrist, Silvia Milano, Beau Madison Mount, Daniel Nolan, Amit Pundik, Sarah Robins, Jon Robson, Sherri Roush, Amia Srinivasan, Margot Strohminger, Alannah Tomich, and Alec Walen for invaluable feedback and discussion. Thanks also to audiences at Birkbeck, King's College London, Oxford University, the University of Tartu, the 2019 Midwest Epistemology Workshop and the 2019 Bled conference on Social Epistemology and the Politics of Knowing Conference. Finally, many thanks to Sherri Roush's graduate seminars at UCLA and the Fall 2019 social epistemology seminar at the University of Tennessee for invaluable discussions.

Bibliography

Allen, Ronald J. (forthcoming) 'Naturalized Epistemology and the Law of Evidence Revisited' *Quaestio Facti*.

Amaya, Amalia (2015) *Tapestry of Reason*. Hart.

Blome-Tillmann, Michael (2015) 'Sensitivity, Causality, and Statistical Evidence in Courts of Law' *Thought* 4(2): 102–112.

Bolinger, Renee (this volume) 'Explaining Justificatory Asymmetries between Statistical and Individualized Evidence' *Truth and Trial*. Zachary Hoskins and Jon Robson (eds.) Routledge.

_____ (2021) '#BelieveWomen and the Ethics of Belief' to appear in *NOMOS LXIII: Truth and Evidence*.

Cohen, Jonathan (1977) *The Probable and the Provable*. OUP.

Di Bello, M. and Collin O'Neil (forthcoming) 'Profile Evidence, Fairness, and the Risks of Mistaken Convictions' *Ethics*.

Enoch, David (forthcoming) 'How to Theorize about Statistical Evidence (and Really, about Everything Else): A Comment on Allen' *Quaestio Facti*.

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

Ferzan, Kimberly Kessler (2021) '#BelieveWomen and the Presumption of Innocence: Clarifying the Questions for Law and Life' to appear in *NOMOS LXIII: Truth and Evidence*.

Gardiner, Georgi (ms) 'She Said, He Said: Rape Accusations and the Preponderance of the Evidence'.

_____ (forthcoming-a) 'Doubt and Disagreement in the #MeToo Era' *Feminist Philosophers on #MeToo* (working title) ed. Yolonda Wilson. Routledge.

_____ (forthcoming-b) 'Legal Evidence and Knowledge' *Routledge Handbook of the Philosophy of Evidence*, eds. Maria Lasonen-Aarnio and Clayton Littlejohn. Routledge.

_____ (2020) 'Profiling and Proof: Are Statistics Safe?' *Philosophy* 95(2): 161–183.

_____ (2019a) 'Legal Epistemology' *Oxford Bibliographies: Philosophy*, ed. Duncan Pritchard. OUP.

_____ (2019b) 'The Reasonable and the Relevant: Legal Standards of Proof' *Philosophy & Public Affairs* 47(3): 288–318.

_____ (2018) 'Legal Burdens of Proof and Statistical Evidence' *Routledge Handbook of Applied Epistemology*, eds. David Coady and James Chase. Routledge, 171–195.
_____ (2015) 'Normalcy and the Contents of Philosophical Judgements' *Inquiry* 58(7): 700–40.
Gersen, Jeannie Suk (2019) 'Assessing Betsy DeVos's Proposed Rules on Title IX and Sexual Assault' *New Yorker*.
Greer, Germaine (2018) *On Rape*. Bloomsbury.
Ho, Hock Lai (2008) *A Philosophy of Evidence Law*. OUP.
Jordan, Jan (2001) *True 'Lies' and False 'Truths': Women, Rape and the Police*. Doctoral thesis, Victoria University of Wellington, New Zealand.
Kelly, Liz, Jo Lovett and Linda Regan (2005) 'A Gap or a Chasm? Attrition in Reported Rape Cases' Home Office Research Study 293, London Metropolitan University.
Koehler, Jonathan (2001) 'When Do Courts Think Base Rate Statistics are Relevant?' *Jurimetrics Journal* 42: 373–402.
Krebs, Christopher, Christine Lindquist, Tara Warner, Bonnie Fisher, Sandra Martin (2007) *The Campus Sexual Assault (CSA) Study*, U.S. Department of Justice.
Lisak, David, Lori Gardinier, Sarah Nicksa, and Ashley Cote. (2010) 'False Allegations of Sexual Assault: An Analysis of Ten Years of Reported Cases' *Violence Against Women* 16(12): 1318–1334.
Littlejohn, Clayton (this volume) 'Justified Belief and Just Conviction' *Truth and Trial*. Zachary Hoskins and Jon Robson (eds.) Routledge.
McCauliff, C.M.A. (1982) 'Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees?' *Vanderbilt Law Review* 35: 1293–1335.
Moss, Sarah (forthcoming) 'Knowledge and Legal Proof' *Oxford Studies in Epistemology* (7). OUP
_____ (2018) *Probabilistic Knowledge*. OUP.
Newman, Sandra (2017) 'What Kind of Person Makes False Rape Accusations?' *Quartz*.
Pardo, Michael (2019) 'The Paradoxes of Legal Proof: A Critical Guide' *Boston University Law Review*.
Pritchard, Duncan (2005) *Epistemic Luck*. OUP.
_____ (2017) 'Legal Risk, Legal Evidence and the Arithmetic of Criminal Justice' *Jurisprudence* 9(1): 108–119.
Redmayne, Mike (2008) 'Exploring the Proof Paradoxes' *Legal Theory* 14: 281–309.
Roush, Sherrilyn (2005) *Tracking Truth* OUP.
Schauer, Frederick (2003) *Profiles, Probabilities, and Stereotypes*. Harvard University Press.
Schmalbeck, Richard (1986) 'The Trouble with Statistical Evidence' *Law and Contemporary Problems* 221-236.
Schoeman, Ferdinand (1987) 'Statistical vs. Direct Evidence' *Noûs* 21(2): 179–198
Simon, Dan (2012) *In Doubt*. Harvard University Press.
Simon, Rita James and Linda Mahan (1971) 'Quantifying Burdens of Proof: A View from the Bench, the Jury, and the Classroom' *Law and Society Review* 5(3): 319–30.
Smith, Martin (this volume) 'Against Legal Probabilism' *Truth and Trial*. Zachary Hoskins and Jon Robson (eds.) Routledge.
_____ (2018) 'When Does Evidence Suffice for Conviction?' *Mind* 127(508): 1193–1218
_____ (2016) *Between Probability and Certainty*: *What Justifies Belief*. OUP.
Thomson, Judith Jarvis (1986) "Liability and Individualized Evidence" *Law and Contemporary Problems* 49(3): 199–219.
Williamson, Timothy (2007) *The Philosophy of Philosophy*. OUP.
Wright, Stephen (ms) 'The Non-Evidential View of Testimony'.
_____ (this volume) 'On Testifying and Giving Evidence' *Truth and Trial*. Zachary Hoskins and Jon Robson (eds.) Routledge.