

Defense Without Threat

Words: 13,687

Abstract

At the heart of the ethics of war and defense is the project of developing a theory of *liability*: a theory of when and why attackers forfeit rights to not be harmed. This essay contributes to this project by developing and defending a heterodox answer to a serious and long-standing challenge to the project — what I call the *Challenge of Merely Apparent Attackers*. I argue that our standard conception of forfeiture is too coarse-grained to adequately answer this challenge, and that we need to distinguish between the forfeiture of one's rights against harm and the forfeiture of the non-constitutive, contingent “perks” of those rights. Appreciating this distinction helps us answer the Challenge of Merely Apparent Attackers as well as a number of other important challenges for the theory of defense.

1 Introduction

Suppose you need a liver transplant. You'll die without one. But the waiting list is too long, and the only way you can get a liver is if you kill your next-door neighbor and steal theirs.

Or suppose you are attacked by a gunman, and the only way to protect your life is to grab an innocent bystander and use them as a human shield, resulting in their death.

It would be wrong to kill either your neighbor or the bystander in these circumstances. It is normally wrong to intentionally harm others, even if by harming them you can prevent comparable harm to yourself. But there are exceptions to this rule. And one of the most well-recognized exceptions is the case of self-defense. Consider a paradigm example:

Paradigm. A villain attempts to murder you. He will succeed unless you kill him first.

You may not steal your neighbor's liver or use a bystander as a human shield to save your life. But you may kill villain to save your life.

The difference is that only Villain has made himself *liable* to be harmed — that is, he has done something to lose some of his normal rights against harm.¹ Your neighbor and the bystander have done no such thing. They retain their normal rights against harm. This is why you are not permitted to harm them. But no such rights stand in the way of harming Villain — hence your permission to harm him.

Liability is *the* central ingredient in most justifications for the use of defensive force.² And so a theory of liability must be at the center of our

¹In the present paper, I'll use the terms 'liable' and 'liability' in a more expansive way than they are sometimes used. Some theorists use the term in such a way that one is not "liable" to some harm when they've lost their right against that harm by way of an act of *consent* (see, for example, Tadros, "Duty and Liability," *Utilitas* 24 (2012): 260). As I'm stipulatively using the term, however, it doesn't matter how one loses the right in question. On my usage, for a person to be liable to ϕ just is for them to lack their normal right against ϕ .

²I say 'most' because some harm impositions are permissible rights-infringements, as when you redirect the trolley away from killing five people and onto a path where it will break one person's foot. Such harm impositions admit of a so-called "lesser-

theory of the ethics of defense. If we want a theory of defense, we need a theory of when and why someone loses rights against harm.

So what is it that makes someone like Villain liable to be harmed? Well it sure seems to have something to do with the fact that *he would harm someone if he isn't harmed him first*. That seems, at first glance, at least a necessary condition for liability.

But there's a serious and long-standing problem with this idea, which is that there are compelling cases where a person seems liable to harm even though they pose no actual threat.³ Suppose someone attempts to kill you, but unbeknownst to either of you their gun is jammed.⁴ Or suppose someone convincingly *pretends* to attempt to kill you, so as to give you the scare of your life.⁵ Neither the "futile attempter" nor the "bluffer" pose any genuine threat of harm. And yet it sure seems they are liable to be harmed.

Cases like these pose a serious challenge for the theory of liability. We need a theory that can account for our judgments in such cases — but that does not at the same time *over-generate* cases of liability. After all, one is not rendered liable merely by *appearing* to pose a wrongful threat. Suppose, for example, that Bloggs receives compelling, but false, testimonial evidence that you're attempting to murder him. You thus appear to him to pose a wrongful threat. Bloggs may even have an evidence-relative justification for harming you. But you plainly are not liable to defensive harm.

evil" justification rather than a liability justification. Many theorists also endorse so-called "agent-relative-prerogative" justifications. See, for example, Jonathan Quong, *The Morality of Defensive Force* (Oxford: Oxford University Press 2020): 58–96.

³See, for example, Alberico Gentili, *De Jure Belli (On the Law of War)*, trans. John C. Rolfe (Clarendon Press 1933)(1598): Bk. I, ch. 14, p. 62–63.

⁴This case is inspired by Helen Frowe's "Apparent Murderer" case from *Defensive Killing* (Oxford 2014): 85.

⁵This case is inspired by Kimberly Kessler Ferzan's case by the same name from "The Bluff: The Power of Insincere Actions," *Legal Theory* 23 (2017):169.

I'll call this challenge the *Challenge of Merely Apparent Attackers*. I think it's a challenge of central importance to the theory of liability, not least because the relevant features of these cases are far from exotic. On the contrary: in the real-world, defensive agents very often operate under conditions of false information. Real-world defensive agents frequently have mistaken beliefs about all sorts of things: about the threat others pose, about what defensive options they have available, and about the potential consequences of those options. As such, an answer to the Challenge is an important part of an ethics of defense that speaks to the conduct of defensive agents in the real world — be they soldiers, police, or private individuals.

As I'll show in this paper, it's also an important challenge because of what we learn about the ethics of defense and the nature of forfeiture in answering it. To date, the most focused attempts to answer this challenge all proceed by positing an additional ground of liability: there's one ground of liability for attackers who pose an actual threat of harm and another ground of liability for merely apparent attackers.⁶ I agree that the challenge shows the need for a more nuanced theory of liability. But I disagree that the nuance concerns the *grounds* of liability. In what follows, I defend an answer to the Challenge of Merely Apparent Attackers where the nuance concerns, not the grounds, but the *properties* of liability. I'll argue for the surprising claim that merely apparent attackers are *not*, in fact, liable to defensive harm, but that we can nonetheless explain why some (and only some) merely apparent attackers bear many of the "upshots" of liability.

On my account, our ordinary notion of rights forfeiture is not quite up to the task of making sense of merely apparent attackers. It's too coarse-

⁶Ferzan, "The Bluff"; Frowe, *Defensive Killing*: 85-86; Renée Jorgensen, "The Moral Grounds of Reasonably Mistaken Self-Defense," *Philosophy and Phenomenological Research* 103 (2021):140-156; McMahan, "Who is Morally Liable to be Killed in War?" *Analysis* 71 (2011): 555-556.

grained. Where actual attackers forfeit certain rights against harm — and all the normal “perks” that come with such rights — the same isn’t true of merely apparent attackers. They do not forfeit rights against harm. But they do forfeit many of the non-constitutive, contingent perks of those rights.

§2 more carefully presents the terms of the challenge. §3 summarizes three existing answers to the challenge, setting the stage for my answer in §4. §5 unpacks important advantages of my account. §6 considers a potential disadvantage of my account. §7 concludes.

2 The Challenge of Merely Apparent Attackers

Consider two cases, one involving a “futile attempter” and the other a “bluffer”:

Jam. Jammy intends to murder Defender. He points a gun at her and prepares to pull the trigger. Defender knows she can prevent Jammy from pulling the trigger only by breaking his leg. Believing that she will be killed otherwise, Defender breaks Jammy’s leg. Unbeknownst to anyone, however, Jammy’s gun is irreparably jammed.

Bluff. Bluffer decides to play a very ill-conceived prank on his workplace manager, Defender. He brings to the office an unloaded gun. During a meeting he stands up, points the gun at Defender, and yells, “Eat lead!”. Defender reasonably believes the gun is real and that she is about to be killed. Knowing she can prevent Jammy from pulling the trigger only by breaking his leg, she does so.

Jammy and Bluffer pose no actual threat to Defender. Defender would not be harmed were she to abstain from harming Jammy or Bluffer. And

yet: Jammy and Bluffer sure seem liable to have their leg broken by Defender. At least they bear the usual *upshots* of liability. Let me explain. When someone has a right not to be harmed, she is typically permitted to defend against threats to that right by force, third-parties are typically permitted to defend against threats to that right by force, and if the right is in fact violated, she typically has the standing to complain and is owed compensation.⁷ We see this clearly in the Paradigm case. Your rights are threatened and you can fight back against Villain, others can fight back on your behalf, and you would be owed compensation and have the standing to complain were Villain to succeed in injuring you.

But now consider the position of Villain — the position of someone who *has* made himself liable to harm. Suppose you fight back against him. Is he permitted to fight back in turn — to engage in counter-defense? Certainly not. Furthermore, a bystander is not permitted to fight back against you on Villain's behalf. And if you do injure Villain in self-defense, you don't owe him any compensation and he has no standing to complain.

So these are four typical upshots of liability. If you are liable to have some harm imposed on you, then it will typically be true that:

(U1): You would not be permitted to fight back against that harm (at least not with comparable or greater force).⁸

⁷At least, defense against a threat to one's rights is typically permissible so long as it is necessary to avert to threat, proportionate in severity, and does not cause too much harm to third-parties.

⁸I insert the parenthetical qualifier in order to leave open the possibility that a person who is liable to be targeted with severe harm *H* may, under some circumstances, be permitted to defend himself with very mild counter-defensive harm. Here's a case. *Reckless loses control of his car while needlessly driving recklessly in poor conditions. The car careens towards Driver. It will kill her unless she throws a grenade that will kill Reckless and redirect his car. Reckless can't redirect his own car, but he can redirect the grenade such that he is unharmed and Defender suffers only a minor injury to one of her fingers.* Here I find it plausible that Reckless is both liable to have the grenade thrown his way *and* that he is liable to defend himself by injuring Defender's fingers. Reckless has at least a lesser-evil justification for harming Defender in this way. But there's also this. Imagine a slightly

(U2): Others would not be permitted to fight back on your behalf (again, at least not with comparable or greater force).

(U3): You would not be owed compensation if you were to suffer that harm.

(U4): You have no standing to complain about that imposition of harm.

Jammy and Bluffer clearly bear all four of these upshots. Neither they nor third-parties would have been permitted to fight back against Defender (they wouldn't have been permitted, for example, to break Defender's leg to prevent her from breaking theirs), they aren't owed compensation for their broken legs, and they lack the standing to complain.

The lesson of the Jam and Bluff cases, then, is that a person can bear these upshots of liability even if they pose no actual threat — even if they are what I'll call a *merely apparent attacker*. We need a theory of liability that explains why.

But we also need a theory that can properly discriminate between cases where merely apparent attackers bear these upshots of liability and cases where they do *not*. Jam and Bluff are two cases where merely apparent attackers bear these upshots. Here are two cases where the merely apparent attacker does not:

False Testimony. Testifier wants to see Innocent harmed by Defender, and so he lies to Defender: he tells her that Innocent

different scenario where Defender has *two* defensive options. *She can either kill Reckless or redirect his car such that he is unharmed but Defender suffers a minor finger injury.* In this case, Defender is required to choose the latter option. (In fact, in this case I think Reckless has a right that Defender choose the latter option, [redacted].) What this shows is that Defender would be required to suffer a minor finger injury to avoid killing Reckless. This fact, I think, contributes to giving *Reckless* a justification for making Defender to bear just such a cost towards just such an end in the first version of the case where it is Reckless's decision whether to redirect the grenade.

is about to kill her. Defender has excellent reason to trust Testifier's testimony, and so she attacks Innocent, breaking his leg.⁹

Evil Twin. While on a road trip, Twin's engine overheats. He walks to the nearest town and enters the first mechanic shop he comes across. Unbeknownst to Twin, however, he has an evil twin brother — a notorious serial killer who has just escaped from prison. Authorities have warned locals that the killer will shoot anyone he comes across on sight. Believing Twin to be the murderer and believing himself to be in imminent danger, the mechanic, Defender, lunges at Twin with a crowbar, breaking his leg.¹⁰

As in Jam and Bluff, the defender in False Testimony and Evil Twin attacks someone who merely appears to pose a wrongful threat. But where Jammy and Bluffer bear the upshots of liability, Testifier and Twin do not. They (and likewise, third-parties) would have been permitted to fight back against Defender, they are owed compensation for their broken legs, and they have the standing to complain.¹¹

So here's the *Challenge of Merely Apparent Attackers*. We need a theory of liability that can capture the intuitive differences between characters like

⁹This case is inspired by Michael Otsuka's "Dignitary" case from "Killing the Innocent in Self-Defense," *Philosophy and Public Affairs* 23 (1994):91.

¹⁰This case is inspired by Jeff McMahan's "Mistaken Resident" case from "Basis of Moral Liability": 387, and Jonathan Quong's "Mistaken Attacker" case from *Defensive Force*: 23.

¹¹This judgment appears to be widely, though not universally, shared. For examples of at least partial endorsement, see Saba Bazargan-Forward, "Defensive Liability Without Culpability," in Christian Coons & Michael Weber (eds.), *The Ethics of Self-Defense* (Oxford University Press 2016): 72; McMahan, "The Basis of Moral Liability": 387; Otsuka, "Killing the Innocent": 91; Quong, *Defensive Force*: 24. In considering a case like Evil Twin, Larry Alexander appears to take the view that Twin would be permitted to fight back, but that a bystander would *not* be permitted to fight back on Twin's behalf. See his "Recipe for a Theory of Self-Defense," in Christian Coons & Michael Weber (eds.), *The Ethics of Self-Defense* (Oxford University Press 2016):29.

Jammy and Bluffer (on the one hand) and characters like Innocent and Twin (on the other). We need a theory that explains why the former bear the aforementioned upshots of liability, but not the latter.

3 Culpability, Responsibility, and Signaling

There have been a handful of attempts to answer the Challenge of Merely Apparent Attackers, all of which proceed by adding an addendum to an account of liability for actual attackers that has the effect of attributing liability to characters like Jammy and Bluffer (but not Innocent and Twin).

One of the most focused and developed attempts comes courtesy of Kimberly Kessler Ferzan. On her view, what makes an *actual* attacker liable to be harmed is that they are *culpable* for the fact that they would otherwise violate someone's right not to be harmed.¹² Culpability also plays a starring role in explaining why merely apparent attackers can be liable to harm. What makes them liable, of course, isn't that they are culpable for posing an *actual* threat of harm (by definition they pose no such threat). Rather, they are liable because they are culpable for *appearing* to pose a threat (i.e., they are culpable for the defender's *belief* that they pose a threat).¹³ We'll call Ferzan's the CULPABILITY ACCOUNT of liability.

CULPABILITY: What makes a person liable to (apparent) defensive harm is that either (i) they are culpable for the fact that someone's rights would otherwise be violated, or (ii) they are culpable for the fact that the defender (reasonably) believes someone's rights would otherwise be violated.

On this account, Jammy and Bluffer bear the upshots of liability because they are in fact liable. And they are liable because they are culpable for

¹²"With respect to the paradigmatic instance of self-defense, an aggressor becomes liable to the defender's force because he has culpably decided to harm the defender and the defender must use force to stop the harm from occurring" (Ferzan, "The Bluff," p. 173).

¹³Ibid, p. 172.

appearing to pose an actual, wrongful threat. Innocent and Twin, by contrast, bear no culpability for their appearing to pose a threat.

Jeff McMahan and (separately) Helen Frowe offer an account that is structurally analogous to Ferzan's, but that replaces the central notion of *culpability* with that of *responsibility*. Call this the RESPONSIBILITY ACCOUNT:

RESPONSIBILITY: What makes a person liable to (apparent) defensive harm is that either (i) they are responsible for the fact that someone's rights would otherwise be violated, or (ii) they are responsible for the fact that the defender (reasonably) believes someone's rights would otherwise be violated.¹⁴

As on Ferzan's account, the McMahan-Frowe account tells us that Jammy and Bluffer bear the upshots of liability because they are in fact liable. And they are liable because they are responsible for appearing to pose an actual, wrongful threat. Innocent and Twin, by contrast, bear no responsibility for their appearing to pose a threat.

Another focused attempt to answer the Challenge comes courtesy of Renée Jorgensen.¹⁵ Jorgensen does not purport to offer a complete theory of liability. Rather, she offers an addendum that is meant to be compatible with many different theories of liability in actual-threat cases. So for Jorgensen, there's the usual conditions for liability that we find satisfied in actual-threat cases — whatever exactly those might be. But then there's a second path to liability. Instead of meeting the usual conditions

¹⁴McMahan, "Who is Morally Liable?": 555-556; Frowe, *Defensive Killing*: 85-86. The above formulation glosses over the fact that, at least for McMahan, it is *comparative* responsibility that matters most. This is the crucial difference between Jammy/Bluffer and Innocent/Twin: Jammy/Bluffer bear more responsibility than Defender for her belief that they pose a threat, whereas Defender bears more responsibility than Innocent/Twin for her mistaken belief about them.

¹⁵Jorgensen, "The Moral Grounds of Reasonably Mistaken Self-Defense," *Philosophy and Phenomenological Research* 103 (2021): 140-156.

for liability, you might make yourself liable by engaging in behavior that conventionally *signals* that you meet the usual conditions for liability. (To illustrate: suppose we attached this addendum to a simple Thomsonian account of liability, according to which an actual attacker is liable when and because he would violate someone else's rights if he were not harmed himself.¹⁶ The result would be that a merely apparent attacker makes himself liable when and because he conventionally signals that he would violate someone else's rights if he were not harmed himself.) I'll call Jorgensen's the SIGNALING ACCOUNT:

SIGNALING: What makes a person liable to (apparent) defensive harm is that either (i) they meet the usual conditions, *C*, for liability (i.e., the conditions specified by the correct theory of liability in paradigmatic, actual-threat cases), or (ii) they have performed an action that conventionally *signals* to others that *p*, where conditions *C* would be satisfied if *p* were true.

Jorgensen thus also holds that Jammy and Bluffer bear the upshots of liability because they are in fact liable. They are liable because they engage in behavior (e.g., pointing a gun at someone, yelling "Eat lead!") that conventionally signals that they pose an actual threat.

We'll come back to these theories later in the paper. For now I mention them as background against which to contrast my own strategy for answering the Challenge. The strategy employed by the above theorists is to posit a second path to liability, alongside the path traveled by actual attackers. Jammy and Bluffer bear the upshots of liability in virtue of traveling this second path. They are not liable for the same reasons that actual attackers are liable, but they are liable nonetheless.

I think there is a great deal to learn from these accounts about the importance of responsibility, culpability, and signaling behaviors. And indeed,

¹⁶Judith Thomson, "Self-Defense," *Philosophy and Public Affairs* 20 (1991): 283-310.

as we'll see, there is much my own answer to the Challenge draws from these accounts. Nonetheless, I think the above answers to the Challenge are mistaken in two respects. First, they are each incomplete: they only tell part of a more complex story. Second, they fail to capture central differences between the moral position of actual attackers as compared to merely apparent attackers.

My own answer to the Challenge takes a very different tack. Rather than proposing a second path to liability, I'll present an account of liability according to which there is only one path. On this account, Jammy and Bluffer are *not* liable to harm, strictly speaking. But the account nonetheless explains why they bear the above four upshots of liability. In §4 I'll unpack this account of liability and the answer it offers to the Challenge. In §5 I'll explain why I prefer this answer.

4 A New Answer

4.1 The Assumption Account

Let me introduce my account by first saying something briefly about the structure of liability. A person is never liable to harm *simpliciter*. Liability is always person-relative and goal-relative.¹⁷ When a person is liable to be harmed, they are liable to be harmed by *so-and-so* as a *means or side-effect of such-and-such goal*. That's the basic structure of the liability relation.

For present purposes, consider this idea that liability is goal-relative. Why think that's true? Well compare:

Reckless Driving 1. Driver is recklessly driving above the speed limit in an effort to make his dinner reservations. He loses control of his car. It hurtles towards Defender. Defender will

¹⁷Here I agree with (among others) Jeff McMahan. McMahan claims that liability is goal-relative in *Killing in War*, (Oxford 2009): 8-9, and that liability is person-relative in "The Limits of Self-Defense," in Christian Coons & Michael Weber (eds.), *The Ethics of Self-Defense*, (Oxford 2016): 201-203.

be killed by the car unless she kills Driver by redirecting his car away from her and into a tree. (Driver will be unharmed otherwise.)

Reckless Driving 2. As before, Driver's car hurtles towards Defender. This time, however, Defender can simply step out of the way, in which case Driver will be unharmed. But Defender also has the option to redirect Driver's car, in which case his body will shield a bystander from the gunfire of a villain in the neighborhood (this will kill Driver).

In either case, killing Driver is a defensive harm; in either case, killing Driver saves the life of someone who is in danger. And yet it is permissible for Defender to kill Driver only in the first case.

The best explanation for this difference is that Driver's liability is specifically tied to the goal of averting *his own threat*. He is liable to be killed in service of *that* goal, but not in service of the goal of averting the villain's attack.¹⁸

[Redacted] I believe the best way to makes sense of the goal-relativity of liability is by appeal to an idea introduced by Victor Tadros.¹⁹ It's the idea that a person can lose certain claim rights by taking on certain kinds of duties. Duties have the right structure to do the work of explaining the goal-relativity feature, since duties are themselves goal-relative. Duties are always duties to *do something*, to *see to some goal*.

The idea, more precisely:

¹⁸I'm not claiming that Driver can never be liable to harm that happens to avert the villain's attack. For example, if Defender's only way to protect *herself* involved redirecting Driver such that he were to shield the bystander, then Driver *would* be liable to be so harmed. But he wouldn't bear this liability in virtue of the fact that the harm is necessary to protect bystander. He would bear this liability in virtue of the fact that the harm is necessary to protect Defender.

¹⁹See his "Duty and Liability" and *The Ends of Harm* (Oxford 2011): 169-196. [Redacted.]

Liability by Duty. When X has not consented to be harmed but is nonetheless liable to be harmed, this is because (i) X has a duty to see to goal G, and (ii) harming X is a means or side-effect of bringing about G.²⁰

Here's how the idea plays out in actual-threat cases like Paradigm. Villain has a basic duty not to harm you. Once he initiates his attack, this duty gives rise to a secondary duty to see that *the costs of his attack not fall on anyone but himself* (since otherwise he would fail a duty not to harm). This is why he is liable to be harmed by *you*: by harming him you are just making him to do what he has a duty to do himself — namely, to see that the costs of his attack not fall on anyone other than himself. His duty to assume the costs of the attack makes him liable to be *made* to assume the costs of the attack.²¹ Call this the ASSUMPTION ACCOUNT of liability.²²

4.2 Assuming the Costs of Another's Attack

Now in a case like Paradigm, Villain's basic duty not to harm is what explains why he takes on a duty to assume the costs of the attack. But a

²⁰I should note that, in making this claim about the explanatory grounds of liability, I don't mean to claim that these conditions are jointly sufficient for liability. My own view departs from Tadros's in that I don't think just any kind of duty can render a person liable to harm that serves the goals of that duty. [Redacted.]

²¹Here's how the Liability by Duty principle helps to explain the liability asymmetry between Reckless Driving 1 and Reckless Driving 2. Driver's duty *not to kill Defender* is much more demanding than his duty to *protect the bystander from being killed by the villain*. Driver's duty towards Defender is so demanding that it would require him even to redirect his car away from her at the cost of his own life (if this were the only way to prevent his killing her). Driver's duty to rescue the bystander is not nearly so demanding. This difference in the demandingness of Driver's own duties, I claim, is what explains why he is liable to be killed in service of the one goal but not the other.

²²This is close to Tadros's own theory of liability to defensive harm. Tadros writes: "One central way of explaining a person's liability to be harmed to avert a threat is to show that she has incurred an enforceable duty to avert that threat, even if she will be harmed" ("Causation, Culpability, and Liability," in Christian Coons & Michael Weber (eds.), *The Ethics of Self-Defense*, (Oxford 2016):116. I agree, but I'm claiming something more general than this. At the heart of the present essay is the idea that a person can have a duty to assume the costs of a violent encounter *without* having a duty to avert some wrongful threat. In fact, as we'll see, a duty of cost assumption sometimes requires the very opposite: it requires a person to *abstain* from averting a wrongful threat.

duty not to harm is not the only thing that can ground a duty to assume the costs of an attack. There are, in fact, a plurality of ways in which a person can incur such a duty. Understanding this is the key to answering the Challenge of Merely Apparent Attackers.

Importantly, we need to notice that there are ways a person can take on a duty to assume the costs of *someone else's* attack. We see this illustrated in a variant of the False Testimony case:

Human Shield. As before, Testifier wants to see Innocent harmed by Defender, and so he lies to Defender: he tells her that Innocent is about to kill her. Defender has excellent reasons to trust Testifier's testimony, and so she attacks Innocent ... As it turns out, however, Testifier can intervene to defend Innocent, but in only one of two ways: by breaking Defender's leg or by shielding Innocent such that Testifier's own leg is broken.

Notice, first, that in this case it is *Defender*, and not Testifier, who attacks Innocent. *She* is the one who most directly causes there to be costs that others must choose how to distribute.

And yet although Testifier isn't the attacker, he is not permitted to let Innocent be harmed by Defender. But neither is he permitted to redirect the costs of the attack back onto Defender — at least not if he has the option to instead redirect those costs onto himself, as he does in this case. After all: Defender is just acting reasonably in light of the false information for which Testifier intentionally gave her. Defender is acting exactly as Testifier foresaw and intended that she would. And Testifier was under no pressure to give her this false information. Given all of these considerations, it seems clear that Testifier has a duty to bear the costs of the attack, rather than let them fall on Innocent or redirect them onto

Defender. He is morally required to shield Innocent from the attack.²³

The important lesson for present purposes is that a person can have duties to assume the costs of someone else's attack. This point is the key to answering the Challenge of Merely Apparent Attackers. After all: merely apparent attackers like Jammy and Bluffer are in much the same position as Testifier. They, too, are highly responsible and highly culpable for inciting Defender to attack someone, where this reaction is reasonable in light of the evidence they have provided her.²⁴ The difference is just that they happen to also be the *target* of the attack. In the Jam and Bluff cases, the person who incites the attack is also the target of the attack. But it's hard to see why that difference should put them under any less of a duty to assume the costs of Defender's attack than is Testifier.

Suppose that's right; suppose Jammy and Bluffer have a duty to assume the costs of Defender's attack. Notice that, on my account of liability, this doesn't establish that they are liable to be attacked by Defender. After all, Jammy and Bluffer don't take on a duty to assume the costs of Defender's attack until *after* that attack has been initiated and there are costs that must fall on someone or other. And of course we cannot appeal to the goal of *that* duty to justify the very actions that bring that duty into existence.

Indeed, this is my heterodox claim: Jammy and Bluffer are not liable to be attacked, strictly speaking, since attacking them does not serve the goal of any duty they have prior to the attack. Jammy and Bluffer have not forfeited rights against harm. *But:* given ASSUMPTION they nonetheless

²³Ferzan, "Provocateurs," *Criminal Law and Philosophy* 7 (2013): 598 sounds a similar judgment: "Iago taunts Othello with false claims that Desdemona is unfaithful. As we know, Othello kills Desdemona. But if Iago had the ability to stop Othello, it seems clear that he should harm Othello or harm himself to save Desdemona. Because his actions provoked Othello's conduct, Iago now not only may, but must, act so as to stop the harm to an innocent."

²⁴Here I agree with Ferzan that culpability matters, and with McMahan and Frowe that responsibility matters. I just don't think it matters in quite the way they claim. More on this in §4.3.

forfeit many of the perks or benefits that normally come with those rights. They bear many of the upshots of liability even though they are not truly liable.²⁵

Consider our four upshots of liability in turn. Start with the *first upshot*: Jammy and Bluffer would not be permitted to fight back against Defender with comparable force. This first upshot follows very straightforwardly from the claim that Jammy and Bluffer have a duty to assume the costs of Defender's attack. Fighting back would be a straightforward way of failing this duty — it would be a way of redirecting the costs of the attack onto Defender.

Turn to the *second upshot*: a bystander would not be permitted to fight back against Defender (with comparable force) on Jammy or Bluffer's behalf. This upshot is less straightforward than the first. It's less straightforward because a bystander, of course, is *not* under a duty to assume the costs of Defender's attack.

But — and here's the key — neither is *Defender* in this case required to assume from Jammy and Bluffer the costs of her attack. That may sound surprising given my claim that Jammy and Bluffer are not liable to be attacked. But reflect for a moment on the Human Shield case from above. Suppose, in that case, that Testifier does what he is supposed to do: he throws himself in front of Innocent as a human shield. As he does so, Defender is made aware of Testifier's lies and Innocent's innocence. And now suppose Defender has a choice. She can allow Testifier to suffer the broken leg in defense of Innocent, or she can redirect the harm onto *herself*. Does Defender have a duty to do this? Does she have a duty to rescue Testifier from a broken leg at the cost of a broken leg to herself?

Surely not. The reason Defender lacks such a duty is that, even though she is the most direct cause of those costs, and even though she is wrongfully

²⁵[Redacted.]

targeting a non-liable person (Innocent), Defender nonetheless bears less responsibility and culpability for those costs than Testifier. In virtue of this difference in responsibility and culpability, Testifier has a duty to assume those costs from Defender, but Defender does not have such a duty towards Testifier.²⁶

Now come back to our characters, Jammy and Bluffer. They are in an analogous position with Testifier in this respect. They bear more responsibility and culpability than Defender for the costs of her attack, and as such their duties of cost assumption are asymmetric.²⁷ Jammy and Bluffer have a duty to assume the costs of the attacker from Defender, but not vice versa.

This, then, is why a bystander may not harm Defender on Jammy and Bluffer's behalf. Defender is liable to be harmed by a bystander only if doing so would make Defender to assume costs she herself has a duty to assume. I've just argued that Defender does *not* have a duty to assume the costs of her attack from Jammy and Bluffer. Thus, she is not liable to be *made*, by bystander, to assume those costs.

Turn, next, to the *third upshot* of liability: Jammy and Bluffer are not owed compensation for their injuries. The grounds of this upshot are much the same as for the second. I've argued that Defender lacks a duty to assume the costs of her attack from Jammy and Bluffer. But a duty to pay compensation for harming Jammy and Bluffer would just be an *ex post* version of the duty to assume the costs of her attack. Paying

²⁶It's important to notice here that duties of cost-assumption are person-relative: you may have a duty to assume the costs of an attack from X but not from Y. In the Human Shield case, for instance, Defender clearly would be required to assume costs from *Innocent*. Suppose Defender could prevent Innocent from being harmed only by assuming the cost of a broken leg: she would be required to do so. But if Testifier throws himself in front of Innocent, Defender does not likewise have a duty to assume the costs of a broken leg from *Testifier*.

²⁷I mention both responsibility and culpability because I they each make a contribution to Jammy and Bluffer's duties to assume costs. See §4.3.

compensation is just what it looks like to assume the costs of an attack after the damage has been done. Since Defender has no duty to assume the costs of her attack, she has no duty to pay compensation after the attack has concluded.

Finally, the *fourth upshot* of liability: Jammy and Bluffer have no standing to complain about Defender's use of force. The standing to complain is — like the permission to use defensive force — neither constitutive of nor necessary for the possession of a right. A person can lose the standing to complain about some treatment without losing their right against that treatment. Demands of reciprocity, for instance, may undercut such a standing. Suppose I'm a pathological promise-breaker; I break almost every promise I make. For this reason, I lack the standing to complain if you break a promise that you make to me. But this isn't because I lack a right — and you the correlative duty — that you keep your promise. (The fact that I don't keep my promises doesn't strip you of your ability to choose to put yourself under a promissory duty towards me.) Rather, I lack the standing to complain in spite of my right. I lack the standing to complain because to complain would be to wrongfully hold you to standards to which I do not hold myself.²⁸

Another way to lose one's standing to complain without losing the relevant right is by knowingly and unnecessarily inciting someone to infringe that right. Suppose I'm a certain kind of provocateur. I know how to *really* get under my colleague's skin. I know that Colleague will be deeply offended and angry if I mock his physical appearance, and that he'll almost certainly respond violently. So I go ahead and slip him a cruel and mocking letter. As predicted, he responds after reading the letter by attacking me. This is something he shouldn't do; it's gratuitous harm that serves no defensive purpose. I may well deserve it, but it's not harm to which I've made myself *liable*. And yet: I plainly have no standing to complain

²⁸[Redacted].

about his attack, in virtue of my role in inciting him to infringe my rights in this way.

But if I have no standing to complain about Colleague's use of force, Jammy and Bluffer surely have even less standing to complain about Defender's use of force. They are just as guilty as I am of inciting violence against them, and Defender has an even stronger excuse than Colleague for her use of violence. Colleague enjoys, at best, a partial excuse for his use of force. Defender is fully excused. So even on the assumption that Jammy and Bluffer have a right that Defender not harm them, we should think they have no standing to complain.

To summarize: characters like Jammy and Bluffer are not liable to be attacked by Defender, in the sense that they simply lack a right against such an attack. This is because attacking them does *not* serve any goal they have a duty to serve. And that's simply because they pose no actual threat. Prior to the defender's reaction, there are no costs that need distributing, and thus no costs that Jammy and Bluffer must be made to assume. This is where my account most drastically comes apart from CULPABILITY and SIGNALING. On the latter accounts, Jammy and Bluffer have forfeited their rights to not be harmed. On my account, they haven't.²⁹ Even detached from the above four upshots of liability, this fact is significant. Jammy and Bluffer's right against harm continues to exert its "deontic force": this right contributes to making it fact-relative impermissible for Defender to attack them.³⁰

²⁹Now although Jammy and Bluffer are not liable *to be attacked*, they do become liable to certain forms of harmful treatment *if and once* they are attacked. Once Defender responds with force, and makes it the case that costs must fall on someone or other, Jammy and Bluffer do become liable to be *made* to assume those costs. So things are like this: Defender wrongs Jammy and Bluffer by attacking them, but does not wrong them by preventing Jammy and Bluffer from redirecting the costs of her attack onto someone else.

³⁰This, I think, is what explains why we could properly blame Defender for attacking Jammy and Bluffer *if* she knew they posed no actual threat. Jammy and Bluffer have a right not to be harmed whether or not Defender is *aware* that they pose no actual threat; either way, Defender lacks an objective justification for harming them. But Defender's

Jammy and Bluffer do not forfeit rights against harm. But they nonetheless forfeit something very important. They forfeit many of the *non-constitutive, contingent perks* that normally come with those rights. They give up the enforceability of those rights, the compensability of those rights, and their standing to complain about the infringement of those rights. They greatly weaken those rights without entirely losing them. To give their status a name, we'll say that they are *quasi-liable*.³¹

ASSUMPTION ACCOUNT. Merely apparent attackers are not liable to defensive harm, strictly speaking. This is because harming them serves no goal that they are under a duty to serve (per LIABILITY BY DUTY). But merely apparent attackers will sometimes be *quasi-liable* to harm — in the sense that they bear some, but only some, of the upshots of liability. When a merely apparent attacker is quasi-liable to harm, this is because they have a duty to assume the costs of someone else's mistaken aggression.

We're finally in a position to see my answer to the Challenge of Merely Apparent Attackers. What distinguishes characters like Jammy and Bluffer from characters like Innocent and Twin is that only the former are quasi-liable, in virtue of the fact that only the former have a duty to assume the

epistemic position does make a difference to her *culpability* for attacking them. When she is not in a position to know that her targets pose no threat, she has an excuse for attacking them. But when she is in a position to know, she lacks such an excuse. She would then *culpably* do what is objectively wrong. More on this in §6.

³¹There are a number of ways in which theorists have previously noted that liability can be "partial". Almost all theorists agree that your liability can be partial in the sense that you are liable only to certain amounts of harm (e.g., only to proportionate and necessary harm). Some theorists think liability can be partial in the sense that you are liable relative to some people and not others (see, for example, David Clark, "Refusing Protection," *Philosophy and Public Affairs* 51 (2022): 33-59; and McMahan, "Limits of Self-Defense": 201-203), or liable relative to some goals and not others (see §4.1, and see McMahan, *Killing in War*, 8-9). What I'm advancing in this paper is the idea that liability can be partial in yet another sense: one can enjoy some, but only some, of the usual perks of having a right.

costs of their defender's attack.

Jammy and Bluffer have such a duty, we've said, because they bear more responsibility and culpability than Defender for the fact that there are costs that must be distributed. But the same is not true of Innocent and Twin. They bear little or no responsibility for that fact. And they are in no way culpable for that fact. That's why they have no duty to assume the costs of Defender's attack, and thus bear none of the upshots of liability.

4.3 Interlude: Responsibility and Culpability

Before I turn to extol the virtues of this story, it might help to address a potential point of confusion. I've claimed to be selling an answer to the Challenge of Merely Apparent Attackers that is importantly different from Ferzan's CULPABILITY account and McMahan and Frowe's RESPONSIBILITY account. And yet the notions of responsibility and culpability play a central role in my answer. Whence the difference?

Three brief comments. First, one point of disagreement has to do with theoretical inclusivity. I think responsibility and culpability are *both* important elements of the full story, and that it's a mistake to have a theory of liability that has a role for one but not the other. A familiar reason to not focus exclusively on culpability comes from cases involving excused attackers — e.g., a case like False Testimony. Defender wrongfully, but non-culpably, attacks Innocent. Nonetheless, Defender's responsibility for the attack puts her under a duty to assume those costs from Innocent, thereby making her liable to defensive harm from Innocent. But there are also reasons to not focus exclusively on responsibility at the expense of culpability. One reason concerns cases where the person who is most responsible for a state of affairs is different from the person who is most culpable for that state of affairs. Here's an example:

Omission. Defender has been given false, but compelling, testimonial evidence that Innocent is about to kill her. Defender

rolls a boulder towards Innocent, believing this is the only way to defend herself. Villain is nearby, and witnesses the scene. He can costlessly divert the boulder such that no one will be harmed. Villain hates Innocent, however, and would be glad to see her killed. He refrains from redirecting the boulder. Innocent, however, is able to defend herself, which she can do either by redirecting the boulder towards Defender or towards Villain.³²

I think Defender is most *responsible* for the threat to Innocent.³³ But it is Villain who is most *culpable* for that threat. And it is clearly on him, rather than Defender, that the costs should fall. This case shows that culpability makes a contribution to liability above and beyond its implications for responsibility. On my view, that's because culpability can contribute to the strength of a person's duties of cost assumption above and beyond its implications for responsibility.

Second, the ASSUMPTION account insists that responsibility and culpability contribute to (quasi-)liability only when and because they produce a duty of cost-assumption. They only work through this intermediary. This may seem like a compatible addendum to RESPONSIBILITY and CULPABILITY. But I don't think it is. This is because — as we'll see in the next section — duties of cost assumption admit of a plurality of grounds other than facts

³²I introduce this case here to note the independent contribution that culpability makes to liability above and beyond the contribution it makes to responsibility. But, as an aside, I'd note that I think liability-by-omission cases like this one generally pose real trouble for Jorgensen's SIGNALING account. Villain makes himself liable to defensive harm, but he performs no behaviors that conventionally signal aggression.

³³Reasons for thinking Defender the more responsible party include the facts that (i) Defender makes a greater causal contribution to the threat to Innocent than Villain; (ii) Defender's contribution to that threat is no less intentional than Villain's contribution-by-omission to that threat; and (iii) Defender is doing harm where Villain is only allowing harm. I take responsibility to be a morally-neutral concept that focuses on the degree of control or "authorship" that an agent exercises over an outcome. Defender exhibits more control over the threat to Innocent than does Villain, and is more aptly described as the "author" of that threat.

about responsibility and culpability. This means that on ASSUMPTION, responsibility and culpability are neither individually necessary nor *jointly* necessary for (quasi-)liability.

Third, as we've already seen, I part ways with RESPONSIBILITY and CULPABILITY in that I think responsibility/culpability for posing an *actual* threat does not have the same kinds of liability effects as responsibility/culpability for *appearing* to pose a threat. When someone poses an actual, wrongful threat, they've already created the costs that need distributing, and their responsibility/culpability for those costs contributes to putting them under a duty to assume the costs. That makes them liable to be harmed in service of that duty. But when someone merely appears to pose a threat, there are not (yet) costs than need distributing. They are thus not (yet) under a duty to assume any costs, and so their responsibility/culpability for appearing to pose a threat does not give rise to any liability. It's only once the defender reacts and generates the need to assume costs that the responsibility/culpability facts generate the duty of cost assumption and the quasi-liability that I've claimed follows.

5 Advantages

The ASSUMPTION account before us, let me now turn to say why I favor the account, apart from the fact that I think it is an upshot of plausible ideas about the relationship between duty and liability and about the ways in which we can take on duties to assume costs.

By far the biggest selling point of the theory is its explanatory reach: it explains a wider range of phenomena than either CULPABILITY or SIGNALING. In this section I'll mention two important respects in which this is so.

5.1 The Challenge of No-Threat Liability

First, ASSUMPTION answers the Challenge of Merely Apparent Attackers while also answering a more general challenge — what I'll call the Chal-

lenge of *No-Threat Liability*.

Here's the thing: cases of merely apparent attackers are just one kind of case where persons who pose no actual threat bear the upshots of liability. There are other cases. The Challenge of Merely Apparent Attackers is just one part of a more general challenge.

I'll give two examples in which a person bears the upshots of liability despite neither posing, nor even *appearing to pose*, any actual threat. The first is an example of what Saba Bazargan-Forward calls "complicitous liability":³⁴

Getaway Driver. Villain wants to see Defender killed. So he hires an assassin to do the job. After sending Assassin on his way, Villain hires someone else to serve as Assassin's getaway driver without Assassin's knowing. The driver waits on the street as Assassin throws a grenade into Defender's home. The grenade will break Defender's leg unless she kicks it towards Driver before it detonates. Defender kicks the grenade towards Driver, breaking his leg.

We can suppose that Driver has not made any causal contribution (even by omission) to the threat posed by Assassin. After all, he makes no difference to Assassin's actions, especially as the latter doesn't even know that Driver is present. Driver knows full well that he is signing up to be a getaway driver for an assassination, but he does not play a role in the planning or execution of the assassination. His role is limited to being ready to contribute *after* the assassination.

Given these stipulations, we can hardly say that Driver is responsible or culpable for the threat to Defender. He is, of course, responsible and

³⁴Bazargan, "Complicitous Liability in War," *Philosophical Studies* 165 (2013): 177-195.

culpable for *something*: he is responsible and culpable for choosing to be Assassin's getaway driver. But he isn't responsible or culpable for the *threat to Defender*, since he makes no contribution to that threat. Nor is he responsible or culpable for the *appearance* of a threat to defender. As such, both the RESPONSIBILITY and CULPABILITY accounts predict that Driver is not liable to have the grenade kicked his way.

We also can hardly say that Driver performs any behavior that conventionally signals aggression towards Defender. He doesn't. And so SIGNALING likewise predicts that Driver is not liable to have the grenade kicked his way.

Driver does not pose a threat. He does not appear to pose a threat. He does not contribute to the (appearance of a) threat that someone else poses. And yet he sure seems liable to be harmed. He bears the upshots of liability. For example: it would be wrong for him to kick the grenade *back* towards Defender. And the reason surely has to do with the fact that he is *complicit* in the assassination, even though he has not contributed to it.³⁵

It's an intuitive idea that complicity can ground liability. One of the virtues of ASSUMPTION is that it helps us to make sense of this idea. This is because it is independently plausible that complicity can ground a duty of cost assumption. We can see this most clearly in the *ex post* context. Imagine, if you will, that Assassin succeeded in breaking Defender's leg and that Driver helps him to escape as planned. Assassin and Driver, however, are later confronted by the police: Assassin dies in a shootout and Driver is taken into custody. Assassin — being dead — is unable to

³⁵Perhaps not everyone will want to describe the relevant relation that Driver bears to the threat as a *complicity* relation. That's fine; I'm not wedded to that particular label. Whatever the best name for the relation, the important point is that Driver's choice to stand ready to aid and abet Assassin in escaping the crime scene relates him to Assassin's threat in such a way as to put Driver under a duty to avert that threat — a duty that is considerably more demanding than the duty a bystander would have to do the same.

compensate Defender for her injuries. Does Driver then have a duty to pay compensation? It sure seems that he does. But if he has a duty to assume the costs *ex post*, it's hard to see why he wouldn't have such a duty *ex ante*.

So ASSUMPTION delivers the intuitive verdict that Driver bears some liability to have the grenade kicked his way. By kicking the grenade his way, Defender is just making him to assume costs he has a duty to assume.³⁶

Here's a second case where ASSUMPTION delivers where competing theories cannot:

Police Shield. A police officer, Patrick, takes an oath upon entering the service — an oath to protect the citizens of his town even at a risk of serious harm to himself. While on patrol one day, he witnesses a gunman enters a public park. The gunman points a gun at an innocent civilian and prepares to shoot at one of her legs. Patrick is not in a position to disarm the gunman, but he is in a position to throw himself between the gunman and the civilian, and take the harm of a broken leg upon himself. Patrick, however, chooses not to shield the civilian. But, as it happens, the civilian has two ways to defend *herself*. She has a metal shield, which she can use either to deflect the bullet away from herself and towards Patrick (breaking his leg) or to deflect the bullet towards another innocent bystander (breaking their leg).

There's a lot we might ask about this case. But I want to focus on one feature in particular, which is that *if* the civilian is going to use the shield, she should deflect the bullet towards Patrick rather than towards the

³⁶Notice that Driver is liable in the strict sense; he's not just quasi-labile. By kicking the grenade his way, Defender is making Driver to assume the costs of an attack that temporally precedes her own use of force.

bystander. There is at least *more* justification for harming Patrick than for harming the bystander.

ASSUMPTION offers an explanation. On this account, Patrick bears some liability to have the bullet directed his way. Not because he poses any threat; not because he appears to pose any threat; not because he is complicit in the attack. Rather, he's liable because he has promised to risk life and limb to defend Civilian. Promises are just another way in which duties of cost assumption might come about. Patrick's promissory duty to assume costs from Civilian explains why there's more reason to deflect the bullet towards him than towards the bystander.³⁷

By contrast, CULPABILITY, RESPONSIBILITY, and SIGNALING don't help to explain this. Patrick is not culpable or responsible for the gunman's threat, and he's performed no signal of aggression. CULPABILITY, RESPONSIBILITY, and SIGNALING thus predict no difference in liability between Patrick and the bystander. And if both persons are equally non-labile, then it's hard to see what other considerations could favor harming Patrick over harming the bystander in precisely the same manner.

The ASSUMPTION account gives us a more complete answer to the general Challenge of No-Threat Liability. The CULPABILITY, RESPONSIBILITY and SIGNALING accounts are instructive. They are instructive in that they help

³⁷Even if Patrick's promissory duty does not make him liable to *that much* harm, his liability would still contribute to a "combined" or "mixed" justification for harming him, which would still explain why Civilian has *more* justification for harming Patrick than the bystander. A combined/mixed justification for some harm is one where part of the harm is justified on the grounds that the target was liable to suffer that much harm, and where the remainder of the harm is justified on lesser-evil grounds. For discussion of such justifications see Bazargan-Forward, "Killing Minimally Responsible Threats," *Ethics* 125 (2014): 114–136; Frowe, *Defensive Killing*, 67-70; McMahan, "What Rights May Be Defended by Means of War?" in *The Morality of Defensive War*, eds. Cécile Fabre and Seth Lazar (Oxford: Oxford University Press, 2014): 133–35; McMahan "Liability, Proportionality, and the Number of Aggressors," in *The Ethics of War: Essays*, eds. Saba Bazargan-Forward and Samuel C. Rickless (Oxford: Oxford University Press): 18-24; McMahan, "Proportionality and Just Cause: A Comment on Kamm," *Journal of Moral Philosophy* 11 (2014): 440; Tadros, *The Ends of Harm*, 253-256.

us to understand *some* of the ways in which duties of assumption can arise. Responsibility and/or culpability for appearing to bear a threat can lead to liability precisely because it can lead to a duty of cost assumption. Likewise, signals of aggression can lead to liability precisely because they too can lead to a duty of cost assumption.

These accounts fall short, however, in that they latch on to only *some* of the ways we can incur such duties. As cases like Getaway Driver and Police Shield illustrate, there are *other* ways. And thus, there are other ways we can become liable to harm.³⁸

5.2 The Challenge of Necessity

Another advantage of ASSUMPTION is that it helps us to make better sense of the so-called “necessity constraint” on defensive harm.

Consider a paradigm case of unnecessary harm in an actual-threat case:

Easy Defense. A villain attempts to murder Defender. Defender knows that she has two equally good ways to disarm Villain: by killing him or by giving him a light slap on the wrist.

It’s widely held that it is wrong for Defender to kill Villain under these circumstances. It’s wrong because it isn’t “necessary” for Defender to kill the attacker to defend herself, and attackers (like Villain) make themselves liable only to necessary defensive harm.³⁹

³⁸Cases involving “hidden attempters” are another kind of case that makes trouble for CULPABILITY, RESPONSIBILITY and SIGNALING. The Two Snipers Redux case in §5.2 is an example of such a case.

³⁹Numerous theorists endorse this idea that the problem with unnecessary defensive harm is that it wrongs the attacker. See, for example, David Clark, “The Demands of Necessity,” *Ethics* (2023); Kaila Draper, “Necessity and Proportionality in Defense,” in *The Ethics of Self-Defense*, eds. Christian Coons and Michael Weber (Oxford: Oxford University Press, 2016): 174-175; McMahan, “Limits of Self-Defense,” 195-197; and Quong, *Defensive Force*: 129-132. But see Frowe, *Defensive Killing*: 88-119, for a dissenting opinion. For a response to (some of) Frowe’s arguments, see McMahan, “Limits of Self-Defense,” 195-206.

This intuitive idea raises an obvious challenge, however. Consider characters like Jammy and Bluffer. They pose no actual threat. And so of course it isn't necessary for Defender to harm them in order to defend herself. The challenge, then, is this: we need to explain why characters like Jammy and Bluffer bear the upshots of liability without undermining our ability to explain why we shouldn't kill Villain in a case like Easy Defense. Call this the *Challenge of Necessity*.

ASSUMPTION has a simple answer to this challenge. On this account, although Jammy and Bluffer bear the upshots of liability, they are not liable to be attacked. They are not liable to be attacked because attacking them does not serve any duty that they're under. It's only *after* the attack has begun that they take on a duty to assume the costs of the attack. Before that, there aren't any costs that need distributing. This is a crucial difference between merely apparent attackers and actual attackers. Actual attackers forfeit rights against harm. Merely apparent attackers like Jammy and Bluffer only forfeit some of the *perks* that come with those rights. What they lose, as we've seen, is the permission that they and others normally have to defend those rights, their claim to compensation for the infringement of those rights, and their standing to complain. They greatly weaken their rights without entirely losing them.

So what ASSUMPTION gives us is an explanation for why Jammy and Bluffer bear the upshots of liability that is perfectly compatible with the principle that attackers are only liable to necessary defensive harm.

By contrast, CULPABILITY, RESPONSIBILITY, and SIGNALING both attribute full-blooded liability to Jammy and Bluffer. And so these accounts are both inconsistent with the idea that liability is restricted to necessary harm. To answer the Challenge of Necessity, then, these accounts need some other principle by which to explain why Defender may not kill Villain in a case like Easy Defense.

Now you might think that it's not so hard to find such a principle. You might think that the problem with killing Villain isn't that it isn't necessary, but rather that it doesn't *appear* necessary. That is, we might explain the case of Easy Defense by appeal to the principle that *liability is restricted to harm that appears necessary*. This reformulation of the necessity constraint is consistent with the thought that Jammy and Bluffer are liable to harm, since harming them does at least appear to be necessary.

But that principle is false, as we can see from cases of hidden attackers. Consider:

Two Snipers. There are two snipers hidden in a belltower, each attempting to kill Defender. Defender only spots Sniper₁, and believes he is the only person in the belltower. She knows that she can save her life only if she throws a grenade into the belltower (she doesn't know that this is also necessary to defend against Sniper₂). She throws the grenade, wounding both snipers.

Sniper₂ is plainly liable to this harm, even though it doesn't appear to be necessary to Defender to harm him. So we should reject the principle that *liability is restricted to harm that appears necessary*.

We might instead try the following weaker, disjunctive principle: *liability is restricted to harm that either is necessary or that appears to the defender to be necessary*. But this principle also runs into trouble. Consider:

Two Snipers Redux. As before, there are two snipers hidden in a belltower, each attempting to kill Defender. Defender only spots Sniper₁, and believes he is the only person in the belltower. She knows that she can save her life only if she disarms him. She can do so by throwing a grenade into one

of two windows. If she throws the grenade into the first window it will break the legs of Sniper₁ and Sniper₂. If she throws the grenade into the second window it will break the legs only of Sniper₁. As Defender is unaware of the presence of Sniper₂, she sees no reason to prefer one window to the other. She chooses to throw the grenade into the first window. Both snipers are wounded. As it turns out, Sniper₂'s gun was jammed and he posed no actual threat.

Sniper₂ surely bears the upshots of liability. He would not be permitted to kick the grenade back towards Defender, bystanders would not be permitted to kick the grenade back on his behalf, and he is not owed compensation for his injuries. Accounts that appeal to full-blooded liability to explain the marks of liability will thus attribute liability here. But notice: it is neither necessary for Defender to harm Sniper₂ (the option to throw the grenade into the other window was available to her), nor does it *appear* to Defender to be necessary to harm him (she doesn't believe that he's a threat; she doesn't even know he's there). The disjunctive principle is also false.

Now perhaps there is yet some way for these accounts to have their cake and eat it too. Perhaps there's some way to insist on strict liability for characters like Jammy and Bluffer without undermining our ability to explain what's wrong with killing Villain in a case like Easy Defense. I don't take myself to have given an exhaustive proof that there isn't a way to strike this balance. There may yet be some alternative principle I've missed that can do the job. What I hope to have shown, though, is that there is a real puzzle here that puts a burden on accounts that attribute full-blooded liability to merely apparent attackers.

It's an attractive feature of ASSUMPTION that it bears no such burden. It doesn't require us to give up the intuitive idea that liability is constrained by considerations of necessity.

And in fact there's another attractive feature of the account here, which is that it helps us to make sense of the fact that it is sometimes permissible to resist unnecessary harm and it sometimes is not. To see this, compare two cases of unnecessary harm:

Irresistible Defense. Villain is fully culpable for attempting to kill Defender. Defender can save her life either by breaking one or both of Villain's leg. But Defender believes she can defend herself only by breaking both of Villain's legs. And so she attempts to break both his legs.

Resistible Defense. Driver is in the mood for pizza. He knowingly risks driving on icy roads to pick one up, and loses control of his vehicle. It careens towards Defender, who is sitting on her front porch. She knows that she can just as well prevent the car from hitting and killing her either by breaking one or both of Driver's legs. But Defender doesn't like Driver very much, and so takes this opportunity to attempt to break both his legs.

In either case, Defender chooses a defensive option that is unnecessary. But there seems an important difference. For one thing, it seems that Driver would be permitted to use at least modest amounts of force to fight back against Defender. Suppose, for example, that he had a way to redistribute the harm such that he and Defender each suffer a broken leg, rather than him suffering two broken legs. I think he would be permitted to do this. But the same is not true of Villain. He would not be permitted to resist in such a way.

We can make sense of this difference by appeal to the idea at the heart of my answer to the Challenge of Merely Apparent Attackers — the idea that a person can have a duty to assume the costs of *someone else's* wrongful attack. Villain is not permitted to fight back because that permission is

undercut by his duty to assume the costs of Defender's mistake — namely, the cost of an extra, unnecessary broken leg. It's independently plausible that Villain has such a duty. And he has such a duty, I think, for reasons of comparative culpability. Defender makes a mistake, yes, but her mistake is at least a reasonable one in light of her evidence. It's a non-culpable mistake. Villain, by contrast, bears a great deal of culpability for putting Defender in a position to make that mistake. Villain's duty to assume the costs of the wrongful harm undercuts the permission he would otherwise have to redirect those costs onto Defender.

Things are different for Driver. He bears minimal culpability for his threat to Defender. And what's more: in his case, Defender *is* at fault for the unnecessary harm she attempts to impose on Driver. She knows better. Under these conditions I find it very plausible that Driver (unlike Villain) is *not* under a duty to assume the costs of Defender's mistake.

That's why Driver can resist and Villain cannot. The difference isn't that the one is liable to unnecessary harm and the other isn't. The difference is that only one has a duty to assume harm that he was not liable to have pressed against him in the first place.

6 An Explanatory Loss?

I've argued that ASSUMPTION has greater explanatory reach than its rivals in that it offers a more complete answer to the Challenge of No-Threat Liability as well as an answer to the Challenge of Necessity. It does better in this respect. But here is one respect in which ASSUMPTION may seem to do worse than its competitors: it lacks a certain unifying promise of its competitors.

To see the worry, consider one of the ways Ferzan motivates her own answer to the Challenge of Merely Apparent Attackers. She does so by appeal to a general principle about our normative powers. Call it the *Can't Fake It* principle:

Can't Fake It. When X culpably causes Y to believe that a normative power has been exercised with respect to Y, X has in fact exercised that normative power.⁴⁰

This principle is plausible when we consider such normative powers as those we exercise by way of consent, abandonment, and promise.⁴¹ "Bluffed" consent has the same normative effect as sincere consent: I can't trick you into committing an act of theft by telling you that you can use my car, even though I don't intend thereby to waive my right that you not use my car. "Bluffed" abandonment has the same normative effect as sincere abandonment: I can't trick you into committing an act of theft by throwing my watch into a public trash can, even though I don't intend thereby to surrender my ownership of the watch. "Bluffed" promises have the same normative effect as sincere promises: when I tell you I'll pick you up from the airport, I put myself under a duty to pick you up from the airport even if I never had any intention of picking you up from the airport or of putting myself under a duty to do so.

According to Ferzan, the same is true of our "power" to *forfeit* our rights. "Bluffed" aggression has the same normative effect as actual aggression. On her account, when Jammy or Bluffer culpably cause Defender to believe that they've forfeited certain rights against harm, they do in fact forfeit those rights against harm. Her account thus treats forfeiture as symmetrical in this respect to the powers we exercise by consent, abandonment, and promise.

My ASSUMPTION account does not. On my account, merely apparent attackers who culpably cause others to believe they've forfeited certain

⁴⁰This is a gloss on what Ferzan ("The Bluff") calls *Forfeiture by Insincere Act*: "If D culpably intends to create in X, or culpably risks creating in X, the impression that he *vs* or intends to *v*, and his actually *v*ing or intending to *v* would result in him lacking a certain right, through that right being forfeited or waived, and X is induced in this way to believe that D lacks the right, then D has in fact forfeited his right against X acting in this way" (192).

⁴¹*Ibid.*, 174-182.

rights against harm do not in fact forfeit those rights. They forfeit many of the contingent upshots of those rights, but not the right itself. The right persists and still provides weighty reasons for the defender not to harm them. ASSUMPTION thus treats “forfeiture” as distinct from the powers we exercise by consent, abandonment, and promise. It places forfeiture outside the scope of the Can’t Fake It principle.

There’s a loss of unity here to be sure. But it’s not a loss of unity to mourn. On the contrary: I think we should want an account that does *not* treat merely apparent attackers as analogous in this respect to those who insincerely consent, abandon, or promise. There are important pretheoretical asymmetries between the former and the latter. Contrast, for example, the following cases:

Insincere Consent. Alice says to Bob, “You can use my car this weekend for your trip from San Diego to San Francisco. You don’t need to return it until Monday.” But Alice doesn’t really mean it: Alice doesn’t intend or want to suspend her right that Bob not use her car; she just wants to be able to accuse Bob of theft in the future. Bob picks up the car on Friday night, and drives to San Francisco. On Saturday — while still in San Francisco — he learns that Alice’s consent was insincere.

Insincere Attack. Alice bluffs that she is going to kill Bob. At first Bob thinks Alice is sincere and poses a threat of lethal harm. But he soon learns that she is lying, that she is holding a paper mâché gun, and that she poses a threat only of giving Bob a mild papercut.

There is an important asymmetry between these cases concerning the normative situation *after* Bob learns about the insincerity. In the consent case, Bob is plainly permitted to continue using the car after learning that Alice’s consent is insincere; he doesn’t need to drive back to San

Diego earlier than planned to avoid committing theft. In the latter case, however, Bob is plainly not permitted to shoot Alice dead after learning that she poses no actual threat.

What does this asymmetry mean for Ferzan's claim that bluffing attackers forfeit rights in the same way that actual attackers do? I see only one way to maintain that claim in the face of the fact that it is obviously wrong for Bob to shoot Alice dead after learning she is bluffing. We would need to say that Alice is liable to be killed *before* Bob knows she is bluffing, but not *after*. And the only way I can see to secure this result would be to make certain claims about the relationship between liability and *proportionality*: we'd need to say that Alice's liability is sensitive to Bob's information in this way because a person is liable only to defensive harm that is *proportionate to the apparent threat*. The thought, then, would be that Alice is at first liable to be killed because such harm is proportionate to the lethal threat she appears to pose, but that once Bob learns about the bluff, lethal harm becomes disproportionate to the very mild threat she appears to pose.

I agree that a person who is liable to harm is only liable to proportionate harm. But the question is how we should understand this requirement. In particular, there is the question of proportionality's relation: *to what* must the severity of defensive harm be proportionate? [Redacted.]⁴² For present purposes I just want to note that the answer cannot be one that appeals to the *apparent* threat. The proportionality constraint cannot be so subjective as this, as we can see by reflecting on a variant of the Two Snipers case:

Bigger Blast. Two snipers are attempting to kill Defender from a nearby belltower. Defender sees both snipers, but she falsely believes that only Sniper₁ poses a threat. She can defend

⁴²[Redacted.]

herself from Sniper₁ only by throwing a grenade into the bell-tower. Defender reasonably but mistakenly believes that the grenade blast will injure only Sniper₁. As it turns out, the grenade blast also injures Sniper₂, breaking one of his legs and averting his attack.

Defender breaks Sniper₂'s leg, imposing harm that is disproportionate to the threat he *appeared* to pose. No matter: he is plainly liable to suffer this harm. There is a proportionality relation that constrains liability, but it is not one that measures defensive harm against the severity of the *apparent* threat.

The lesson is that treating forfeiture as a normative power subject to the Can't Fake It principle gets us into really hot water. This gain in theoretical unity comes at too great a cost. It requires us to adopt implausibly subjectivist accounts of the proportionality constraint on liability (just as it requires us to adopt implausibly subjectivist accounts of the necessity constraint, as argued in §5.2).

I agree that it is plausible that those who insincerely consent, abandon, and promise have the same normative standing as those who sincerely do these things. But we've seen that there is motivation, independent of my account, to think that merely apparent attackers do not have the same normative standing as actual attackers. I consider it an advantage, not a cost, of ASSUMPTION that it captures and illuminates this difference.

7 Conclusion

Let's take stock of where we've come. There's a challenge for the theory of liability — the Challenge of Merely Apparent Attackers. I've argued that the ASSUMPTION account offers the resources to answer this challenge. ASSUMPTION claims that a person is liable to be made to bear harm they haven't consented to when and because they'd be required to take that

harm upon themselves. On this account, merely apparent attackers like Jammy and Bluffer are not liable to harm, but they nonetheless bear many of the upshots of liability. They are *quasi-liable*.

Moreover, I've argued that this answer to this challenge also helps us answer the more general Challenge of No-Threat Liability as well as the Challenge of Necessity. It even helps us to understand why some unnecessary defensive harms can be resisted and why some can't be.

A closing note. As I've presented it here, ASSUMPTION is incomplete. I've claimed that duties of assumption are the grounds of liability and quasi-liability. I've claimed that there are a variety of ways such duties might arise, and I've pointed to a few such ways. For example, I've claimed that a person can have a duty to assume costs in virtue of their duty not to harm, or in virtue of their responsibility for those costs, their culpability for those costs, their complicity in those costs, or because they have made certain promises. But this is where there is a great deal more to be said. This is a research project that can't be tackled in a single paper. Filling out ASSUMPTION requires cataloguing these different grounds of assumptive duties and of understanding how they interact with one another.

What I hope to have accomplished in the present paper, however, is to motivate the idea that *this* where much important action lies in the theory of defense. I hope to have shown that the way to make progress in understanding liability across the full range of cases is to make progress in understanding when and why we have duties to assume costs from one another.

References

- Larry Alexander, "Recipe for a Theory of Self-Defense," in Christian Coons and Michael Weber (eds.), *The Ethics of Self-Defense* (Oxford University Press 2016).
- Saba Bazargan-Forward, "Complicitous Liability in War," *Philosophical Studies* 165 (2013): 177-195.
- , "Killing Minimally Responsible Threats," *Ethics* 125 (2014): 114-136.
- , "Defensive Liability Without Culpability," in Christian Coons and Michael Weber (eds.), *The Ethics of Self-Defense* (Oxford University Press 2016).
- David Clark, "Refusing Protection," *Philosophy & Public Affairs* 51 (2022): 33-59.
- , "The Demands of Necessity," *Ethics* 133 (2023): 473-496.
- Kaila Draper, "Necessity and Proportionality in Defense," in *The Ethics of Self-Defense*, eds. Christian Coons and Michael Weber (Oxford: Oxford University Press, 2016).
- Kimberly Kessler Ferzan, "The Bluff: The Power of Insincere Actions," *Legal Theory* 23 (2017): 168-202.
- , "Provocateurs," *Criminal Law and Philosophy* 7 (2013): 597-622.
- Helen Frowe, *Defensive Killing* (Oxford: Oxford University Press, 2014).
- Alberico Gentili, *De Jure Belli (On the Law of War)*, trans. John C. Rolfe (Clarendon Press 1933)(1598).
- Renée Jorgensen, "The Moral Grounds of Reasonably Mistaken Self-Defense," *Philosophy and Phenomenological Research* 103 (2021): 140-156.
- Jeff McMahan, "The Basis of Moral Liability to Defensive Killing," *Philosophical Issues* 15 (2005): 386-405.
- , *Killing in War* (Oxford: Oxford University Press, 2009).
- , "Who is Morally Liable to be Killed in War?" *Analysis* 71 (2011): 544-559.

- , “What Rights May be Defended by Means of War?” *The Morality of Defensive War*, eds. Cécile Fabre and Seth Lazar (Oxford: Oxford University Press, 2014).
- , “The Limits of Self-Defense,” in *The Ethics of Self-Defense*, eds. Christian Coons and Michael Weber (Oxford: Oxford University Press, 2016).
- , “Liability, Proportionality, and the Number of Aggressors”, in *The Ethics of War*, eds. Saba Bazargan and Samuel Rickless (New York: Oxford University Press, 2017).
- , “Proportionality and Just Cause: A Comment on Kamm,” *Journal of Moral Philosophy* 11 (2014): 428-453.
- Michael Otsuka, “Killing the Innocent in Self-Defense,” *Philosophy and Public Affairs* 23 (1994):74-94.
- Jonathan Quong, *The Morality of Defensive Force* (Oxford: Oxford University Press, 2020).
- Victor Tadros, *The Ends of Harm: The Moral Foundations of Criminal Law* (Oxford: Oxford University Press, 2011).
- , “Duty and Liability,” *Utilitas* 24 (2012): 259-277.
- , “Causation, Culpability, and Liability,” in *The Ethics of Self-Defense*, eds. Christian Coons and Michael Weber (Oxford: Oxford University Press, 2016).
- Judith Jarvis Thomson, “Self-Defense,” *Philosophy and Public Affairs* 20 (1991): 283-310.