



Liability, culpability, and luck

Dana Kay Nelkin¹ 

Accepted: 18 January 2021 / Published online: 18 February 2021
© The Author(s) 2021

Abstract This paper focuses on the role of culpability in determining the degree of liability to defensive harm, and asks whether there are any restrictions on when culpability is relevant to liability. A natural first suggestion is that it is only relevant (or at least is significantly enhanced) when combined with an actual threat of harm in the situation in which defensive harm becomes salient as a means of protection. The paper begins by considering the question of whether two people are equally liable to defensive harm in a situation if both culpably intend to harm another, but due to circumstances outside the control of the two people only one has a chance of succeeding in causing harm. I argue that there is no difference in liability between the two. I then turn to a kind of slippery slope challenge that accepting this conclusion would lead to a vast over-inclusiveness in those liable to defensive harm, and consider a recent attempt at meeting it that requires that a person's culpability can only affect liability if it concerns the very situation in which defensive harm is relevant. Finally, I put forward and assess a new way of meeting the challenge that appeals to a particular conception of culpability together with auxiliary theses concerning how culpability can decrease over time, among others.

Keywords Culpability · Liability · Responsibility · Luck · Defensive harm · Ethics of war

✉ Dana Kay Nelkin
dnelkin@ucsd.edu

¹ Department of Philosophy, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0119, USA

1 Introduction

Questions about what makes one liable to harm arise in a number of distinct literatures where relevant harms can take a number of forms, including punishment, blame, harm that results from self- or other-defense, and harm in the course of warfare. In each literature, the distinction between the culpable (or blameworthy) and non-culpable looms large. In this paper, my focus will be on liability to defensive harm and its relationship to culpability, though I will draw on insights from other discussions of liability to harm. In turn, I aim for the conclusions to inform some of the other debates, most notably those involving the ethics of war, where the parallel to defensive harm is frequently explored in great depth.¹

2 Culpability and a challenging set of cases

Consider a pair of cases:

Culpable Threat: A person acting in a fully culpable way is chasing you with a loaded gun and will kill you unless, at a precise moment, you pull a lever that you know will cause a huge boulder to fall onto the person chasing you, thereby killing him.

Bystander: A person acting in a fully culpable way is chasing you with a loaded gun and will kill you unless, at a precise moment, you pull a lever you know will cause a huge boulder to fall in the person's path, but will fall directly onto a non-culpable bystander, thereby killing him.

The only difference in the two cases is the identity of the person below the boulder when you can release it, and the difference between them is that one culpably created the threat to you and one did not. Does the culpability of the person under the boulder make a difference as to how you are permitted to act? It is intuitively compelling that culpability makes a difference, and it seems that the reason that it makes a difference is because it affects the threat's liability. A person is liable to the imposition of harm, as I will understand liability here, when it would not violate her rights or in any way wrong her to impose such harm.²

¹ See, for example, Frowe (2014) who describes and endorses the rise of what she calls "reductive individualism," according to which "it is a mistake to think of war as a morally distinctive enterprise, governed by rules that are irreducible to the moral rules of ordinary life." In contrast, on this view, "the rules that govern harming in war are the rules that govern harming between individuals. Defensive war is not, on this view, *another* exception to our prohibition on the use of force, in addition to the exception that permits individual defence. Rather, war is the very *same* exception as individual defence" (p. 2).

² I take this to identify the broadest and most inclusive concept of liability. A particular *conception* of liability might then be one that requires responsible or culpable intentional action to disarm the right. (For the distinction between a concept, which every party to the debate takes themselves to be trying to offer an account of, and particular conceptions of the concept about which parties disagree, see Rawls (1971). It is not always made explicit in the literature whether a debate is occurring at the level of concept or conception. For example, Frowe writes that, "If I say that Attacker is liable to a defensive harm, this is just a concise way of expressing the following claim: 'Attacker lacks a right against having the defensive harm inflicted upon him, because he has forfeited his usual rights against such harm as a result of

The difference in culpability could make *all* the difference when it comes to permissibility, such that you are simply permitted to release the boulder in Culpable Threat and not permitted to do so in the Bystander case. If we just focus on a binary choice as is given in the cases, it might seem that the liability of the person under the boulder is all or nothing. But consistent with this verdict in the particular pair of cases is the idea that what one is liable for comes in degrees, where this is reflected in our being able to impose one of a range of harms on one person that is greater than that which one is permitted to impose on the other (Tadros 2016; Arneson 2018). We can further test the idea that there is differential liability by imagining a combined case in which there are two levers, one that will cause a boulder to fall onto the culpable person chasing you with the loaded gun and one that will cause a boulder to fall onto the non-culpable bystander. It seems clear that releasing the boulder over the threat creator should be the preferred choice to save oneself.³

The general thesis of interest here is the following:

Culpability Matters: Whether and/or to what degree an agent is culpable in a given situation is relevant to their degree of liability to defensive harm.

As it stands, the thesis is somewhat vague, and can be made more precise along a variety of dimensions. And importantly, there is no commitment that culpability is either necessary or sufficient for the permissibility of defensive harm.⁴ But even in this form, it nicely explains our intuitive reactions to the cases described above.

Despite the intuitive appeal of the idea that culpability makes a difference to agents' liability and thereby to whether or what degree of harm you are permitted to impose, some have offered powerful arguments against it, while in some

Footnote 2 continued

behaving in a particular way' (2014, p. 72). Frowe goes on to argue that one can only be liable to defensive harm if one is morally responsible for posing an unjust threat. As a claim about the precise conditions for liability that in principle competes with others (such as McMahan's (2009) claim that liability depends on the responsible loss of a right, or Tadros's (2012) claim that liability depends on one's having an enforceable duty), it is a particular conception of liability. But the claim that liability requires forfeiture seems also to be a claim about which there can be disagreement, and this suggests that either there are overlapping concepts of liability and the disagreement is merely apparent, or that this itself is part of a particular conception of a single and very general concept of liability. For the purposes of this paper, which is focused on the relationship between culpability and liability, I believe that it is not necessary to resolve this last question. Thanks to a reviewer for pressing me to situate my understanding of liability within the larger debate.

³ It is not obvious exactly what the relationship is between being liable for more defensive harm and being ahead in the queue for defensive harm. For it seems open that two people could meet a threshold of liability to defensive harm—say, being infected with a cold in order for one to save one's life—and yet one person is liable for no more than that while another is liable for death. In that case, one wouldn't wrong either by infecting them with a cold to save oneself, but in the comparative case it seems that the one liable for more should be first in line. This suggests that liability generated by culpability on the part of an individual alone is only part of the story whether people ought to impose that which one is liable for. And yet the amount of liability appears to appropriately affect our judgments about who is first in line. Compare Tadros's (2016) "Selection Principle", and Arneson's approach in (2019).

⁴ Note that this principle has a parallel when it comes to punishment. Such a position is consistent with the denial that punishment is non-instrumentally good, as well as with the position that culpability provides no reason on its own that the culpable be punished or receive what they deserve.

notable cases acknowledging its serious intuitive appeal. (See, for example, Thomson 1991; McMahan 2005; Quong 2015). While compelling replies have been given to many of those arguments, I believe that one challenge to the idea that culpability matters remains inadequately addressed. It can be represented as taking different forms, but one such form is as a slippery slope argument: if culpability matters in the cases above, then it seems that we will be led inexorably to a commitment to its making a difference in a whole host of other cases where it does not intuitively seem to matter. (See McMahan 2005).

To see this, consider the following cases:

Culpable Attempter: A person acting in a fully culpable way is chasing you with a loaded gun and will kill you unless, at a precise moment, you pull a lever that you know will cause a huge boulder to fall in the agent's path, but onto a second culpable agent, who is also trying to kill you with a gun but whose gun you alone know is not loaded. If the boulder falls onto the second agent, it will kill him.

Disposed Threat: A person acting in a fully culpable way is chasing you with a loaded gun and will kill you unless, at a precise moment, you pull a lever that you know will cause a huge boulder to fall in the agent's path, but onto a person watching the scene who is disposed to kill you and would if he could, but who believes correctly that his gun is not loaded. If the boulder falls onto the second person who is disposed to kill you, it will kill him.

Glee Taker: Same as *Disposed Threat*, except that the person is not disposed to try to kill you. At the same time, he takes great glee in anticipating your imminent death.

Yesterday's Threat: A person acting in a fully culpable way is chasing you with a loaded gun and will kill you unless, at a precise moment, you pull a lever that you know will cause a huge boulder to fall in the agent's path, but onto a second person who is not carrying anything, but who yesterday tried just as hard and culpably to kill you with a loaded gun before the gun jammed. If the boulder falls onto the person who tried to kill you yesterday, it will kill him.

Last Week's Threat to Someone Else: Same as Yesterday's Threat, except that the person did not threaten you, but instead culpably threatened someone else with some degree of harm last week.

Lifelong Non-Culpable Agent: Same as Bystander, except that the person has never done anything wrong throughout his entire life, let alone culpably.

It is counterintuitive in at least several of these cases that there is liability for defensive harm. So, ideally, defenders of Culpability Matters will be able to draw a principled line separating these cases into two groups, showing why agents in some cases are liable to defensive harm on account of culpability and in some cases not. If apt, they might also draw multiple lines distinguishing between different degrees of liability in various cases. In doing so, they will be required to make the thesis of Culpability Matters more precise.

Presumably, if Culpability Matters, then at least Culpable Threat and Lifelong Non-culpable Agent should be treated differentially when it comes to defensive harm on account of a difference of liability generated by their difference in culpability. I do not mean to suggest that the cases are already properly ordered so as to accept a single line between some consecutive pair, and as we will see there is dispute about just how the cases should be ordered.

In the remainder of this paper, I assess three candidates for line-drawing. In Sect. 3, I begin with an influential proposal, according to which culpability for successfully causing a threat generates greater liability than culpability without causation. On this proposal, a line is drawn between Culpable Threat and Culpable Attempter. There are two versions of this proposal. The first draws a solid line between them, arguing that culpability *and* prospective causation are necessary conditions for liability, so that Culpable Attempter is not liable to defensive harm at all. The second draws a softer—or “dotted”—line, arguing that both are liable, but that Culpable Threat is *liable to more* defensive harm than Culpable Attempter, and that Culpable Threat is closer to the front of the queue for defensive harm. I assess two different arguments for drawing a line (continuous or dotted) between the original Culpable Threat and the Culpable Attempter, and tentatively conclude that neither argument is sound. If this is right, then we are left with the challenge of showing where a line (or lines) *should* be drawn. In Sect. 4, I assess an initially appealing alternative suggestion that limits the relevance of culpability in a given situation to those cases in which a person is culpable “with respect to the situation” *involving* the threat even if they are not the cause of the threat. In Sect. 5, I then propose and assess a distinctive account that builds in more restrictions to Culpability Matters by, in part, appealing to a particular conception of culpability, namely one that appeals to the opportunity to avoid wrongdoing.

Before I turn to the line-drawing suggestions, it is important to explain how I understand culpability. As I will understand it, one is culpable for an action or omission just in case one is responsible and blameworthy for it. In particular, one is blameworthy in the sense that one is accountable for it.⁵ A question immediately arises whether this notion is the same as, or different from, the notion of desert. The concept of accountability is clearly different from the concept of desert, although I believe that the conditions for instantiating each are the same. I have argued for that mutual entailment in detail elsewhere, though here I hope as much as possible to abstract from those details. One reason that has been given for setting desert to the side is the view that an agent’s being deserving of a harm entails that it is a non-

⁵ I take it that this is the “accountability” sense of blameworthiness, distinguished by Watson (1996) from the “attributability” sense, which is an “aretaic” sense—meaning that it is appropriate to attribute the manifestation of a vice, or at least a bad end, to the agent in question. On my view, attributability does not entail accountability; one might exhibit a vice while being very young or having a serious psychiatric disorder that undermines the idea that one is accountable despite being rightly described as having acted for bad ends or with viciousness. I take it that ‘accountability’ is a concept for which there are a variety of competing conceptions. I return to questions about the conditions for being accountable in Sect. 5.

instrumental good that the agent receives that harm and thereby a *pro tanto* reason to impose it simply because it is deserved.⁶ But as I have also argued elsewhere, one's being deserving does not entail that it is such a good, though being deserving can, together with other background conditions, be part of a reason for imposing it. Thus, I think it is a mistake to think that desert is an *entirely* different topic, but for current purposes I will set it aside as much as possible.

3 The first suggestion: luck counts, or drawing a line between Culpable Threat and Culpable Attempter

An influential suggestion for drawing a line and avoiding a slippery slope is that culpability matters to a greater extent when combined with causation of the threat. In other words, those who are culpable for *causing* the threat at hand are liable for harm, and those who are not culpable for causing the threat at hand only because their efforts failed⁷ are either not liable or liable for less than their successful counterparts. More precisely,

More Culpable for Causing Threats, More Liable: In aiming to undo or prevent a serious evil or injustice, in a case when one person must be harmed, agents are first in the queue based on greater liability for defensive harm when they are seriously culpable for causing the threat at hand (even if others are seriously culpable but have not caused the threat).⁸

Note that this is not intended to be a complete answer to the challenge of the Slippery Slope Argument. In its current form (representing a dotted line), it might be thought to slow, without stopping, the slide. Nevertheless, it is a partial answer insofar as it suggests that nothing short of culpably causing the threat at hand can generate the *fullest* level of liability.

3.1 An argument by analogy to the role of luck in compensation and punishment

The “moral outcome luck” debate begins with a question about whether two people who form identical intentions to commit a crime are equally blameworthy, despite the fact that the world happens to intervene and prevent the crime from being carried out in one case and not in another. For example, suppose two responsible

⁶ See McMahan (2005, p. 386) who sets desert aside, but then speaks of culpability without addressing the relationship between the two. On my view, being deserving of a harm itself yields a kind of liability to harm, without itself providing any reason to impose such harm. The same conditions the satisfaction of which makes one culpable also make one deserving. See Pereboom (2014) for a discussion of an important notion of responsibility in terms of basic desert.

⁷ I would also include those who would have been culpably negligent in risking the creation of a threat but get “lucky” in that the threat never materializes. But for ease of discussion, I focus on the cases with which we began, which include intentional actions.

⁸ See Arneson (2018) for the language of “aiming to undo or prevent a serious evil or injustice.”

agents attempt murder with equal wholeheartedness and skill, but only one succeeds due to the fact that one of the prospective victims stoops to pick up a penny at the precise moment the bullet is flying just above. It is a widely held (but by no means universally accepted) view among moral and legal theorists that the two agents are equally blameworthy on the grounds that when it comes to what was in their control they did the same things. The only difference between them was a matter of luck. Thus, those who take them to be equally blameworthy deny what has come to be known as “moral outcome luck” for blameworthiness. However, even among those who deny such luck for blameworthiness, many take it that there *is* moral luck for liability *when it comes to compensation and punishment*.⁹ The one who actually harms another is on the hook for more in the way of either or both of compensation and punishment. After all, what is there for the mere attempter to compensate for? If this is right, then, by analogy, it seems that we ought to do the same when it comes to liability to defensive harm.¹⁰

Let us explore a particular line of reasoning in favor of differential liability in such a case when it comes to punishment, as I think it can help us see why—in the end—it does not help support a line between Culpable Threat and Culpable Attempter when it comes to liability to defensive harm. In an influential article on punishment, David Lewis argues that there is nothing unfair about a criminal justice system that treats attempted crimes like lotteries; if one happens to succeed and cause a lot of damage, then one draws the short end of the stick, so to speak. One took one’s chances, and now one gets the greater punishment, and there is nothing unfair in this (see Lewis 1989). If this is correct, then it is but a small step to the conclusion that liability for an attempt is less than the liability of one in possession of equal culpability who also succeeds in causing harm.¹¹ (Note that this argument grants the denial of moral luck when it comes to blameworthiness and culpability; the conclusion is that liability to compensation and punishment can vary depending on luck *even if* there is no moral outcome luck for blameworthiness).

This is an appealing line of reasoning. But notice that when it comes to differential punishment, we are not so much asking whether one is liable to more punishment than another, but whether it would be unfair or not to punish differentially. Thus, both the successful murderer and the mere attempter could in fact be equally liable, while there is nothing unfair about, say, letting the mere attempter off with less than his liability would make fair. A plausible diagnosis of

⁹ See Khoury (2018) for a recent example of someone who argues against the existence of moral outcome luck for blameworthiness, while nevertheless leaving open a kind of luck for liability to compensation.

¹⁰ Note that this style of argument is suggested by McMahan (2005, pp. 403–404) even though he does not accept a Culpability Matters account. He does recognize that such an account will overlap with his own favored Responsibility Account, however, and so sees that even on his view there is a need for a serious answer to a similar kind of line-drawing challenge. I note, however, that his treatment of Culpability views assumes that they take culpability to be both necessary and sufficient for liability, and neither Culpability Matters nor the more precise versions to be examined make such a claim of either necessity or sufficiency.

¹¹ Technically, it is “would-be” causation—causation if defensive action is not taken. See McMahan (2005), among others, for a view that requires causation for liability, in the sense that the person is an actual threat, or in other words, would cause death (or other harm) without defensive action.

what is going on here is that when it comes to punishment, we are really interested in institutional distributive *fairness*, where fair systems may be consistent with a variety of views about individual pre- or even post-institutional liability (Nelkin 2013). Suppose, for example, that in another possible world, deterrence was better and more efficiently achieved by punishing mere attempters and successful murderers equally. As long as they were not punished more than they were liable for and there was no other way to achieve the ends of a peaceful and stable society, it does not seem that it would be unfair to do so. This intuition, which for me is quite strong, suggests that liability to punishment could be the same in both cases. This argument by analogy to punishment does not succeed, then, in supporting the idea of differential liability.

And yet, it must be conceded that, at least at first, it is more intuitive that only a person who actually causes harm is liable for *compensation*. And liability for compensation does appear to be the closer cousin to liability for defensive harm and so is the better case as an analogy supporting the drawing of a line between Culpable Threat and Culpable Attempter.

Yet there is good reason to think that when we look at a larger set of reflective intuitions about liability for compensation, they do not point toward any sort of line between these cases after all. To see this, imagine a system of tort law in which those who are culpably negligent and thereby cause bad outcomes and those who are culpably negligent and luckily avoid bad outcomes are all asked to pay equally into a fund for compensation. To give this system a coherent rationale, we can suppose that identification of negligent offenders, even when lucky in not causing bad outcomes, is close to perfect. And we can further suppose that in order for the fund to remain solvent at all times, it needs to have on hand an amount equivalent to all negligent actors paying an equivalent of the harm caused or the harm that they would have caused had their negligence caused harm. It seems to me that such a system would not be unfair. What can we conclude from this? As we saw in the case of punishment, the justice of such a system of compensation, precisely by virtue of its being a *system* or institutional structure, answers a different question than that of individual liability. But just systems are not equally consistent with *any* theory of individual liability. The fact that such a system would not be unfair seems to me to suggest that individuals who are culpably negligent, even if they are lucky, are liable for paying the same compensation (or incurring the same cost) as the negligent who cause harm simply as a result of their negligence. Here, the intuition, for me, is that such a system would not even require those who succeed to “catch a break” and pay less than they are liable for so that they can match the maximum level of their unsuccessful counterparts.

It might be objected that once we are in the realm of the non-criminal law, a whole host of systems of payment are fair without regard to liability *or* culpability. A system that taxes everyone seems fair, regardless of culpability. But as luck egalitarians have pointed out—and one needn't be a luck egalitarian to see the point—it can strike us as unfair in a way if the non-culpable pay specifically for the culpable and serious bad choices of a few culpable people. We might think no one is entirely off the hook, we all make some bad choices and these come out in the wash, so to speak. And it is plausible that many things people have been thought in the

past to be culpable for that are the most costly for society are simply not things for which people generally are culpable—e.g., treatment for addiction, smoking-related disease, and so on. Thus, it might be that little revision of the relevant aspects of current practice would be needed to accommodate this point. But to the extent that people *are* culpable for very costly outcomes (e.g., rescues of risk-taking snowboarders), it can seem unfair to ask everyone to chip in an equal share to cover such choices. And yet it seems not unfair to ask the lucky culpable to contribute.

Thinking about these systems allows us to rebut the argument at hand by seeing that there may be no difference between the pre-institutional liability of the lucky and unlucky when it comes to both compensation and punishment. At this point, to the extent that one shares the intuition about the fair system of compensation that equalizes demands for successful and unsuccessful attempters without giving succeeders a lucky break, we can now go even further by adopting the very same analogical structure of the argument in favor of the opposite conclusion. If liability for defensive harm is analogous to liability for compensation in this way, then there is no difference between the lucky and the unlucky when it comes to liability to defensive harm either.¹²

Thus, while this argument has a great deal of initial attraction, reflection on the strategy undermines its appeal, and actually suggests a positive argument to the contrary.

Finally, it might be objected that the advocate of this argument is mistaken in conceding to the opponent the denial of moral luck for blameworthiness, and it is this step that allows the opponent to employ the same analogical strategy in support of the opposite conclusion. For if the advocate of the original conclusion instead *accepts* moral outcome luck for blameworthiness, taking the murderer to be more blameworthy than the mere attempter, then there is a much simpler path to the conclusion. To see this, we can note that someone who violently assaults another is more culpable than someone who by negligence causes another to suffer a paper cut. Plausibly, more culpability leads to more liability to punishment, and also to liability to defensive harm as well. Thus, if there is moral outcome luck when it comes to culpability itself, the argument for differential liability between Culpable Threat and Culpable Attempter becomes much easier to defend.

A full reply would include an elaboration of the many arguments against moral outcome luck, a task I cannot complete here.¹³ For now, I note that it is far less intuitive that Culpable Threat and Culpable Attempter are differentially culpable than that the assaulter and the negligent paper-cut-causer are. Beyond this, my reply here to the argument by analogy will have to remain conditional on the denial of moral outcome luck for blameworthiness. At the same time, I believe that this discussion of the dialectic brings out something of interest: on either view, causation is not an *additional* consideration above and beyond culpability. Either it is already built in as a *culpability* enhancer, or, on the view I favor, it is irrelevant.

¹² It is important to note, however, that one might reject this further positive argument and nevertheless accept the negative one undermining the analogical reasoning in favor of drawing the line between the successful and unsuccessful attempters.

¹³ See Nelkin (2019a) for an overview of the debate about moral luck.

In the remainder of this section, I focus on a second argument, recently developed by Victor Tadros.

3.2 Tadros's doing and allowing argument

Tadros begins by offering a case that combines something like Culpable Threat with something like Culpable Attempter, to help us test whether a causal condition might play an important role after all by asking us to order a queue of targets of defensive action. It is important to note, however, that Tadros does not think of liability as all or nothing, but instead offers an argument for a dotted line between the two kinds of cases, claiming that culpability and causation are both relevant.

First, consider his case:

Unread Letter: Veronica and Wilma each wish to kill Dan. Independently, each sends a letter to Kev, a hit man. Each letter instructs Kev to kill Dan with a pistol at noon. Kev receives Veronica's letter. Wilma's letter gets lost in the mail. Kev immediately acts on Veronica's instructions, finds Dan, and attempts to kill him at noon. Veronica, Wilma, and Irene, an innocent, are standing by. Had Dan received Wilma's letter rather than Veronica's, he would have acted in exactly the same way. (Tadros 2016, p. 119)

Now suppose, with Tadros, "that Dan can save his life only by forcing either Veronica or Wilma to prevent Kev from killing him. Whoever is forced to do this will be harmed to degree n ... Let us suppose that n is sufficiently small that Veronica would be liable to be harmed to this degree [were Dan's harming her the only way to save his life]. This will be so even if n is very large" (p. 120). Whom should Dan harm in saving his life? Tadros suggests that it is intuitive that he ought to select Veronica, and that, if that is so, then a causation condition plays a role in determining degree of liability. If this is right, it would mark a distinction between Culpable Threat and Culpable Attempter.

The argument so far rests on eliciting an intuition that I, for one, do not have. But Tadros has a further argument against those who would resist drawing such a line and take Veronica and Wilma to be equally liable:

Suppose that the *Doctrine of Doing and Allowing* (DDA) is true: we have more stringent duty not to kill than not to let die. Now suppose that Kev kills Dan. Either Veronica or Wilma could have prevented this by jumping in front of the bullet that Kev fires at Dan. Veronica has killed Dan. Wilma has let Dan die. Veronica has more stringent duty to protect Dan than Wilma, for she has more stringent duty not to kill than not to let die, and if she does not jump in front of Dan, she, unlike Wilma, will have killed rather than let die.

Now, consider things from Dan's perspective. If Dan uses Veronica, he forces her to do something that she has a stringent duty to do. If Dan uses Wilma, he does to her something that she has less stringent duty to do. This consideration militates powerfully in favor of Dan forcing Veronica rather than Wilma to avert the threat. (2016, p. 121)

As we saw before, if that is correct, then there is strong reason to draw at least a dotted line between them. Further, Tadros notes, one might resist this argument by rejecting the DDA, but he claims that it “rests on the powerful idea that we are especially responsible for the things that we do, and less responsible for the things that others do” (2016, p. 121). I also wish to preserve the DDA, but I do resist the argument.

The key to my reply is a different understanding of the best version of the DDA.¹⁴ According to Tadros the DDA says that “we have more stringent duty not to kill than not to let die.” The best version of the DDA, in contrast, allows that under *some* circumstances, our duty not to let die and our duty not to kill might be equally stringent. Consider the pair of cases used by James Rachels to try to cast doubt on the DDA. Smith will gain a large inheritance from his uncle if his 6-year old cousin drowns in the bathtub. Smith goes into the bathroom while his cousin is taking a bath, and pushes him under the water until he drowns. Jones is in the same situation, and enters the bathroom with the same intention, but he sees that his cousin has already slid under the water. Though Jones could rescue him, he is delighted that he does not have to exert any effort to accomplish his goal, and stands by as the cousin dies (Rachels 1975). Smith and Jones seem equally horrible, and to have committed wrongs of equal magnitude. Rachels concludes that there is no morally relevant difference between killing and letting die (and it is possible to conclude more generally that there is no relevant difference between doing and allowing). But this needn’t be the end of the DDA. For, as Foot (1977) points out, there might indeed be a relevant difference that is not captured by the bottom line judgments in this pair of cases. She concludes that, all else equal, killing is more difficult to justify than letting die. But in some cases, they might both be equally unjustifiable. And it is possible that the same thing can be said about Wilma and Veronica. Alternatively, we can take the DDA to simply state that there is a morally significant difference between doing and allowing harm, such that it sometimes makes a bottom-line difference to the stringency of our duties not to do and allow harm, favoring allowing over harming, when keeping all else fixed.¹⁵ For example, Kagan (1988), in a similar spirit to Foot, allows for this possibility, suggesting that the problem with so-called “contrast cases” like Rachels’ is that they mistakenly presuppose that moral considerations simply contribute independently to an overall sum of moral justification. There might instead be interactive effects, so that in some otherwise similar cases of doing harm and allowing harm—depending, say, on motive, or other obligations—the duty not to allow harm might be just as stringent as the duty not to harm. (See also Quinn 1989, pp. 289–290.) Thus, by adopting a

¹⁴ One might also question the idea that Veronica would “kill” Dan if he does not harm someone, given that she writes a letter on which Kev acts. But I set this aside as the example might be changed to eliminate the additional agent.

¹⁵ See Quinn (1989) for an account that recognizes a morally significant distinction on the basis of there being different kinds of reasons against harming and allowing harm, and Rickless (1997) for an account in the spirit of Foot’s but that also formulates the principle itself in terms of there being a morally significant distinction without the language of one being in all cases “harder to justify.” Woollard (2015) develops an account in the spirit of Quinn’s, while, like Foot, construing the DDA in terms of doing harm being “harder to justify” than allowing harm.

more subtle version of the DDA, we have room to see the wrongs in his case as equally bad, and their consequent culpability as equally strong.

At this point, however, one might argue that though the two cases strike us at first as equally *bad*, a more subtle test can show that even in the bathtub cases there is just the sort of bottom-line difference that Tadros identifies between the violation of a more stringent duty and the violation of a less stringent duty. Frances Kamm suggests that a test “is to see if it would be permissible to impose the same losses on Jones and Smith, if these losses were to bring their victims back to life” (2007, p. 17). In Kamm’s view, it would not: “Although it might be permissible to kill Smith, it would not be permissible to kill Jones.”¹⁶

Interestingly, this intuition does not appear to be universally shared, even among defenders of the DDA.¹⁷ Here is another kind of test that can be applied to Unread Letter that appears to point in the opposite direction: suppose that none of Veronica, Wilma, or Dan know whose letter Kev received and whose got lost in the mail. Does Dan have an obligation to find out before choosing one to block Kev’s shot? It does not seem to me that Dan would be doing anything wrong by not checking first (assuming he had a few seconds in which to do so).

I imagine, however, that not all will share this intuition, and we might ultimately have a stalemate of intuitions about both Unread Letter and the cases meant to help us decide between versions of the DDA. If that turns out to be our situation, it does not appear that Tadros’s appeal to a particular conception of the DDA will help resolve the dispute, since adopting that conception itself depends on the very same sorts of intuitions that are unshared. The test appealed to for stringency of duty is precisely a test of intuitions for liability to harm, and yet, this is what is at issue.

Further, it is notable that the argument’s premises yield an arguably unattractive implication. For if degree of blameworthiness is partly a function of the severity of the wrong, which is itself a very plausible thesis, then it follows that Veronica is more blameworthy than Wilma. This is to accept moral outcome luck for blameworthiness, a thesis which is quite controversial. It is initially, at least, *more* plausible that Veronica and Wilma are differentially liable (which is what Tadros

¹⁶ In defending the doing/allowing distinction in the face of arguments by Michael Tooley that are similar to Rachels’, Frowe (2007) uses a test of what one may do in self-defense to someone who will otherwise kill one and to someone who will otherwise let one die. While she does not apply it to the particular Bathtub cases (or to Tooley’s that are most similar to these), the test has been taken to be one that would show a bottom-line difference in moral status in all otherwise similar pairs of cases. See Woollard and Howard-Snyder (2016).

¹⁷ Note, too, that even if one has differential bottom-line intuitions in the Bathtub cases, this is not enough to show that all otherwise similar pairs of cases must have bottom-line differences in terms of stringency of duties. Make the case one in which the cousins are replaced by parents or sworn caregivers. In this case, it is plausible that the duty not to allow harm is just as stringent as the duty not to harm. Or, to take a pair of cases of Kagan’s, consider two cases of self-defense, one of which involves pushing one’s attacker into a well and one of which involves not warning him that he is about to fall into it. As Kagan suggests, both seem equally justified, and this suggests that the fact that one is a doing and one an allowing does not make a bottom-line difference in moral status. Even the existence of one such case opens up the possibility that Veronica and Wilma could be equally liable to harm while the DDA can yet be preserved.

claims explicitly) than that they are differentially blameworthy (which he does not claim). And yet the argument's premises yield both conclusions.

Thus, while the argument is intriguing, I believe it does not in the end prompt a revision of intuitions for those who do not already have them. Perhaps in the end, the conclusion ought to be accepted, and some intuitions revised. But I believe that we should only decide whether to do this after seeking worthy alternatives for responding to the Slippery Slope Argument. And it is important that both arguments in some way fare better on the assumption that there is moral outcome luck; and yet, if this assumption is accepted, then the condition of causation is not independent of culpability after all.

4 A second suggestion: drawing the line at involvement in the situation

According to Arneson, culpability-related liability is limited only to culpability *with respect to the situation* at hand:

Fault Forfeits First (Situation version): In aiming to undo or prevent a serious evil or injustice, in a case when one person must be harmed, it is morally required that one choose an agent who is seriously culpable with respect to that evil or injustice and more culpable than anyone else involved in the situation. (Arneson 2018, p. 235 and 260)¹⁸

Thus, on his view, Culpable Threat and Culpable Attempter can be treated in the same way, and tied for the front of the queue, assuming equal culpability.¹⁹ But Fault Forfeits First (Situation Version) offers an expansive condition on which Disposed Threat and Glee Taker will also all be liable to defensive harm in the situation described, while Yesterday's Threat and Last Week's Threat to Someone Else will not be.²⁰ Arneson takes as advantages that the account stops the counterintuitive slide all the way down the slope, and that it captures intuitions.

But as Arneson acknowledges, it is not clear what the rationale is for the restriction to the situation. The restriction does, in part, track a feature of the criminal law—namely, that one is to be judged for one's behavior in the situation in which charges have arisen, and not for past crimes. And while that might have excellent grounding in the purposes of the criminal law and justified protections against misuse, it is not obvious that it is irrelevant in other circumstances. If your partner forgets to pick up your child once, forgiveness might be the order of the day; if he forgets to pick up your child ten times in a row, it is justifiable to react with less

¹⁸ Note that the label, "Fault Forfeits First", is taken from Feinberg (1970).

¹⁹ It must be noted that Arneson does not explicitly claim that there is a tie for first in the queue between Culpable Threat and Culpable Attempter. But applying his own conception of culpability as a matter of conscientious taking into account others' rights, it would seem that they are equally culpable in the situation.

²⁰ Though this latter point is not strictly implied by the statement of the condition, Arneson makes clear elsewhere that he does not take the agents in these cases to be liable, and this is one way we can see that the slide down the slope is stopped.

than forgiveness on this last occasion. Thus, when it comes to these cases, the condition is potentially underinclusive, or at least lacks a rationale for its being restrictive in this way.

One might wonder, in addition, whether the account is overinclusive, particularly if one focuses on the glee-taker. Is the mere having an attitude sufficient to make one liable (even to a degree) for defensive harm? This is surely a less intuitive case than that of Culpable Threat. Note, however, that although Arneson describes the glee-taker as culpable for his glee, it is often disputed that people are responsible and blameworthy for their emotions, which are not under their direct control. Thus, identifying the right conception of culpability is essential for figuring out where to draw lines for liability, and I explore this point further in the next section.

5 A third suggestion: employing a higher threshold for present culpability

In light of both the insights and the difficulties of the suggestions we have so far examined, I propose that we proceed even further down the slope for the moment in a way that is nevertheless principled:

Fault Forfeits First (Non-situation-restricted version): In aiming to undo or prevent a serious evil or injustice, in a case when one person must be harmed, agents are first in the queue based on liability for defensive harm when they are seriously culpable and more culpable than anyone else involved in the situation.

This account takes culpability to be relevant, but with no restriction to cases of causation of threat or to cases in which the culpability is “with respect to the situation.” Now at first glance, one might think that we have rolled all the way down the slope after all. But there are several resources available to the defender of this version of Fault Forfeits First to respond to this worry. The first set is to combine this suggestion with a particular conception of culpability that is both plausible *and* that presents a relatively high bar for inclusion.

On the view I favor, one is culpable to the extent that one acts wrongly with a high enough quality opportunity. Having an opportunity of sufficient quality is in turn a function of both the normative competence, or the capacity to respond to reasons on the part of the agent, together with the congeniality of the situation. For example, the quality of one’s opportunity might be undermined by either or both of situational constraints (e.g., duress) or one’s own reasons-responsive capacities (one’s mathematical skills are sub-par and yet needed to disarm a deadly bomb in time to prevent it exploding in a crowded market). If it is easy to act well and one does not take the opportunity to do so, then one is more culpable, other things equal, than if it is very hard to do and one fails to do the hard but right thing. This means that one is not culpable when one lacks an opportunity (or even when one lacks a high enough quality opportunity). Three important further features of the account must be noted. First, one need not have a good opportunity to avoid wrongdoing at the time of the wrongdoing, as long as one had a good enough opportunity earlier to

have avoided being in the current situation in which one cannot avoid doing wrong.²¹ Second, having the relevant opportunity, in my view, requires having some sort of awareness of risk. It might be only a very general kind of awareness, so that one is unaware of the details of a potential future harm, but without awareness of any kind, one simply lacks a good enough opportunity to avoid wrongdoing. And finally, one can have a high quality opportunity to avoid wrongdoing even if one doesn't act intentionally or exercise one's agency at all. One might simply be aware of a potential risk, have the opportunity to avert it, and make no decision at all about whether to act or not.²²

This view is consistent with treating attempts and successes similarly.²³ And yet, some of the cases described simply will not count as an agent having had such an opportunity and failing to take it. An agent who is merely disposed to culpable wrongdoing, for example, will not thereby have had (let alone failed to take) an opportunity to act.²⁴ Similarly for the glee-taker, particularly a glee-taker who is not even disposed to wrongdoing. This is not to say that such agents are not somewhere in the queue for targets of defensive harm, but on the Quality of Opportunity view, they are not in the queue for reasons of culpability and liability.²⁵

By appealing to the Quality of Opportunity conception of culpability, we are able to deal with some of the initially troubling cases by categorizing them as non-culpable after all. Here, the particular conception of culpability becomes crucial; a conception according to which one is culpable just insofar as one displays a bad quality of will, say, will have a much larger extension, and this point can help explain why Arneson, in applying his Culpability in the Situation account, does so in an expansive way. But raising the bar for culpability in the way that the Quality of Opportunity view does makes the current suggestion significantly less expansive than it would otherwise be.

²¹ This kind of condition is often called a "tracing" condition. See Fischer and Tognazzini (2009) for a classic discussion.

²² For further elaboration of these points, see Nelkin and Rickless (2017). Frowe (2014), in arguing that moral responsibility generates liability in the case of indirect threats, appeals to the idea that "a person is morally responsible for posing an unjust threat if she had a reasonable opportunity to avoid posing the unjust threat and she intentionally failed to avail herself of the opportunity" (p. 73). Being morally responsible for posing the unjust threat in turn renders one liable for defensive harm, and in such a case, defensive harm is permissible. Further, on her view, one may kill an indirectly threatening person only if she is morally responsible for posing an unjust threat (p. 86). Frowe's condition on moral responsibility differs from the one set out in the text in some key ways that are worth noting: the condition is in one way more restrictive than that in the text in that it only includes intentional choice, but it is less restrictive in that it is only presented as a sufficient and not a necessary condition. In contrast, the account in the text is drawn from a general and complete account of moral responsibility and has the resources to give definitive verdicts in a wide variety of cases of defensive harm, including ones without intentional action.

²³ See Nelkin (2019b) for this sort of view applied to moral luck.

²⁴ If, on the other hand, the agent is culpable for having cultivated the disposition, then the situation may be somewhat different. Such a variant of the case may be assimilable to cases like Yesterday's Threat or Last Week's Threat to Someone Else. See below for discussion of such cases.

²⁵ Arguably, a world in which someone disposed to harm is himself incapacitated is a better world in which someone not so disposed to harm is incapacitated. And this might give one more reason to bring about the first world than the second *if* one were going to bring one of the about. But that is not to say that one ought to bring either about.

Even with this important addition, the account is still an expansive one, insofar as there is no restriction to the situation, let alone causation of the threat. And this poses a serious challenge for the view: Do we think that someone who culpably committed a serious crime ten years ago ought to be liable to be kicked off the shoulder of a road where he otherwise has a right to stand now? Someone who had broken a lunch date? Where would we stop?

The suggestion is not without further resources, however. Suppose the person who committed a crime ten years ago fully repented and has done all sorts of things to discharge duties she incurred as the result of her wrongdoing. Then we might say that, despite having committed the crime, her liability is erased at this point, and, arguably because she is now less *blameworthy*.²⁶ This allows the defender of this suggestion to resist at least some counterintuitive implications in cases in which people who committed serious transgressions a long time ago are first in the queue even now.

Yet another resource for defending More Culpable for Causing Threats, More Liable is the requirement that the culpability be serious, where seriousness is best understood as a function of degree of responsibility and severity of the wrong. Appeal to this condition staves off worries about people becoming vulnerable to serious defensive harm just in virtue of having culpably and unjustifiably broken a lunch date. Of course, the question arises what justifies this condition. Here it seems to me that there is a kind of proportionality constraint, just as in the case of punishment.²⁷

Still, one might ask, even with all three of these ways that the scope of this version of Fault Forfeits First is decreased, whether there is nothing about culpability in the actual situation that matters? For example, if ten years ago someone committed a transgression with high culpability, and now one faces a threat by another whose culpability in causing it is only marginally lower, and one can save oneself only by harming one of these two people, should the one who committed the transgressions ten years ago be ahead in the queue?

It is somewhat difficult to ensure that all the relevant stipulations of the case are in fact kept fixed so that we do think that the person whose transgression was ten years ago is just as culpable as when it was committed.²⁸ But suppose we can ensure this and the intuition remains. It is still open to the defender of this Fault Forfeits First version to accommodate it by saying that *liability* is one thing, and all things considered permissibility is another. Perhaps there are additional norms that govern preferred targets when multiple candidates exceed the bar for liability to defensive harm. But another option is to adopt an insight from the “Situation” account, and yet instead of *restricting* liability in that way, instead take it simply that the situation

²⁶ See Khoury and Matheson (2018) for an appealing argument that blameworthiness can decrease depending on what one does.

²⁷ In addition, as a contingent matter, almost all of us are culpable for less serious transgressions, and given that, it seems that one would not be in a position to assume that someone else in the situation was more seriously culpable, or culpable for more than oneself or another potential victim.

²⁸ See Ryazanov et al. (2018) for the difficulty in ensuring that people accept the stipulations in thought experiments, especially if they appear unrealistic.

can *matter*. The idea is that *either* the causation of the threat or the attempt to cause a threat *in the situation* can also enhance one's liability relative to another who is equally culpable, but not with respect to the situation. This suggestion inherits one of the concerns of the restricted version, namely, the absence of a rationale, although the rationale is now for a weaker claim. Additional interesting questions remain about just how *much* a factor relative to culpability this should be. For example, suppose that due to mild negligence on the part of one person, another person finds herself under serious threat of harm. She can save herself by inflicting serious harm on the cause of the current threat or on someone else who earlier in the morning culpably tried to kill her (even if success was never a possibility because, say, there were no bullets in the gun's chamber). Abstracting from the question of whether in either case harming is all things considered permissible, and focusing just on the question of who leads the queue based on liability, intuitions in this case pull at least somewhat in the other direction. But I must leave these here as questions for further exploration. Importantly, the defender of this version of Fault Forfeits First can add further ways of drawing lines.

Leaving open the addition of a situation-related enhancing condition, I tentatively conclude that the Quality of Opportunity account of Culpability, combined with a view of Culpability Matters and auxiliary claims about the conditions under which one can, over time, shed the culpability that underlies liability, offers a promising package that is worthy of further consideration.

6 Conclusion

I have defended the claim that there is no difference in liability to defensive harm between Culpable Threat and Culpable Attempter, while offering what I take to be a promising principle for drawing lines between these cases and some of the other cases that pose the slippery slope challenge such as Glee Taker. At the same time, this suggestion, like the others considered, is not intended to be a complete categorization of all cases, but is meant to capture an important line.

Before concluding, it is worth emphasizing that the package I have proposed is itself consistent with an even larger package of conditions for permissibility. For example, it is not committed to liability being *restricted* to cases of culpability. McMahan's claims about mere responsibility being sufficient could be appealing as an additional condition, and at least *some* non-culpable choices are surely ones in which one voluntarily assumes liability and is thereby liable (McMahan 2005). And beyond liability, there may be times when one is permitted to act defensively against a completely innocent and non-liable person. Finally, it is important to note that this view does not take a stand on exactly what conditions would have to be to make defensive harm permissible. It does however, commit to serious culpability always being a factor, even if outweighed or cancelled out by others, and to offering a framework for how culpability partially determines the ordering of a queue for permissible defensive harm.

Acknowledgements Many thanks to Helen Frowe and Massimo Renzo, organizers of the Conversations on War workshops on moral responsibility (co-hosted by the Stockholm Centre for the Ethics of War and Peace, Stockholm University and the Yeoh Tiong Lay Centre for Politics, Philosophy and Law, King's College, London), for the opportunity to think about the issues raised in the paper and their relation to the ethics of war, and to them and all of the participants, including Helen Beebe, Yitzhak Benbaji, Gunnar Björnsson, Romy Eskens, Jimmy Goodrich, Alex Kaiserman, Ben Matheson, Carolina Sartorio, Derk Pereboom, and Victor Tadros, for their very helpful input and feedback. I am also very grateful for stimulating discussions of these topics with Richard Arneson and Sam Rickless, and to two anonymous reviewers for valuable comments on previous drafts.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arneson, R. (2018). Self-defense and culpability: Fault forfeits first. *San Diego Law Review*, *55*, 231–263.
- Feinberg, J. (1970). Sua culpa. In J. Feinberg (Ed.), *Doing and deserving: Essays in the theory of responsibility* (pp. 187–221). Princeton: Princeton University Press.
- Fischer, J. M., & Tognazzini, N. A. (2009). The Truth about Tracing. *Nous*, *43*, 531–556.
- Foot, P. (1977). Euthanasia. *Philosophy and Public Affairs*, *6*, 85–112.
- Frowe, H. (2007). Killing John to save Mary: A defense of the moral distinction between killing and letting die. In J. K. Campbell, M. O'Rourke, & H. S. Silverstein (Eds.), *Action, ethics and responsibility* (pp. 47–66). Cambridge: MIT Press.
- Frowe, H. (2014). *Defensive killing*. Oxford: Oxford University Press.
- Kagan, S. (1988). The additive fallacy. *Ethics*, *99*, 5–31.
- Kamm, F. M. (2007). *Intricate ethics: Rights, responsibilities, and permissible harm*. Oxford: Oxford University Press.
- Khoury, A. (2018). The objects of moral responsibility. *Philosophical Studies*, *175*, 1357–1381.
- Khoury, A., & Matheson, B. (2018). Is blameworthiness forever? *Journal of the American Philosophical Association*, *4*, 204–224.
- Lewis, D. (1989). The punishment that leaves something to chance. *Philosophy and Public Affairs*, *18*(1), 53–67.
- McMahan, J. (2005). The basis of moral liability to defensive killing. *Philosophical Issues*, *15*, 386–405.
- Nelkin, D. K. (2013). Desert, fairness, and resentment. *Philosophical Explorations*, *16*, 117–132.
- Nelkin, D. K. (2019). Thinking outside the (traditional) boxes of moral luck. *Midwest Studies in Philosophy*, *43*, 7–23.
- Nelkin, D. K. (2019a). Moral luck. In E.N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2019 edition), <https://plato.stanford.edu/archives/sum2019/entries/moral-luck/>.
- Nelkin, D. K., & Rickless, S. C. (2017). Moral responsibility for unwitting omissions: A new tracing account. In D. K. Nelkin & S. C. Rickless (Eds.), *The ethics and law of omissions* (pp. 106–129). Oxford: Oxford University Press.
- Pereboom, D. (2014). *Free will, agency, and meaning in life*. Oxford: Oxford University Press.
- Quinn, W. (1989). Actions, intentions, and consequences: The doctrine of doing and allowing. *The Philosophical Review*, *98*, 287–312.
- Quong, J. (2015). Proportionality, liability, and defensive harm. *Philosophy and Public Affairs*, *43*, 144–173.
- Rachels, J. (1975). Active and passive euthanasia. *New England Journal of Medicine*, *292*(2), 78–80.

- Rickless, S. C. (1997). The doctrine of doing and allowing. *The Philosophical Review*, 106, 555–575.
- Ryazanov, A., Knutzen, J., Rickless, S. C., Christenfeld, N., & Nelkin, D. K. (2018). Intuitive probabilities and the limits of moral imagination. *Cognitive Science*, 42, 38–68.
- Tadros, V. (2012). Duty and Liability. *Utilitas*, 24, 259–277.
- Tadros, V. (2016) Causation, culpability, and liability. In: C. Coons and M. Weber (Eds.), *The ethics of self-defense*.
- Thomson, J. J. (1991). Self-defense. *Philosophy and Public Affairs*, 20, 283–310.
- Watson, G. (1996). Two faces of responsibility. *Philosophical Topics*, 24, 227–248.
- Woollard, F. (2015). *Doing and allowing harm*. Oxford: Oxford University Press.
- Woollard, F., & Howard-Snyder, F. (2016). Doing vs. Allowing harm. In: E.N. Zalta (Ed.) *The Stanford encyclopedia of philosophy* (Winter 2016 Edition). <https://plato.stanford.edu/archives/win2016/entries/doing-allowing/>.

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