

## CHAPTER 8

### *The conditions of liability to preventive attack*

*Jeff McMahan*

The objection to preventive war that is perhaps the most compelling is also, unsurprisingly, the most common. It is that to the extent that the prevention of future aggression is accepted as a just cause for war, the constraint against the initiation of war will be correspondingly eroded. For the claim that another state will, unless prevented, engage in aggression at some point in the future can be cynically exploited as a pretext or public rationale for virtually any unjust war of aggression. To the extent that states would avail themselves of this pretext for the resort to war, they would also become more fearful of becoming the target of an allegedly preventive war. Given the strategic advantages of striking first, each state in an adversarial relation with another would then have an increased incentive to strike first just to avoid being the victim of a first strike.<sup>1</sup> Wars for which the public justification is that they are preventive thus tend to decrease security everywhere.

The reason this objection is so compelling is that it suggests that the acceptance of a doctrine of preventive war could have dreadful consequences, particularly in areas plagued by settled animosities, such as those between India and Pakistan, North and South Korea, and Israel and most or all of its neighbors. This objection, in other words, gives us reasons, as individuals, to *fear* any tendency to recognize preventive war as a form of just war.

In this essay, I will focus on a different concern about preventive war, one that is less a matter of prudence, or practical concern, and more a matter of moral principle. It is a concern that I have explored in two previous essays, so my aim here will be to try to take the discussion deeper

<sup>1</sup> See David Luban, "Preventive War," *Philosophy and Public Affairs* 32:3 (2004), 207–248. The logic of this objection to preventive war parallels that of the main objection in the literature on nuclear strategy to the development of a "counterforce capability."

than I was able to in that earlier work.<sup>2</sup> The question I will address is this: if preventive war can be justified at all, what *form* might the justification take? Because preventive war, like other forms of war, involves intentionally killing people, a moral justification for preventive war must be a moral justification for intentionally killing people. While there is, I will suggest, a form of justification that explains the permissibility of killing in at least some wars of national self-defense against an actual attack, there are reasons to doubt whether this form of justification can be applied to preventive wars.

There are at least four possible forms of justification for intentionally killing a person. Retributivists believe that it can be permissible – or, as Kant thought, morally required – to kill a person who *deserves* to be killed. If a person deserves to be killed, assuming that is possible, he must have acted in a way that not only stripped him of his right not to be killed but also gave others, though perhaps only those with proper authority, a moral reason to kill him that is not instrumental in character. That is, the reason is not that the killing is a means to a further end. According to one interpretation, what this means is that, while the killing is bad *for* the person who deserves it, it is nevertheless intrinsically or impersonally good.

Another possible justification for killing a person is that the person has freely consented to be killed, or waived her right not to be killed. Normally, a person's free and informed consent to be killed is on its own insufficient for the permissibility of killing her. There must also be an independent and substantial reason to kill her. The most common such reason is that it would be objectively better for her to die than to continue to live. Thus many people, myself included, believe that it can be permissible to kill a person if it would be bad *for her* to continue to live *and* if she freely consents (or better yet, forcefully pleads) to be killed.

I will assume that no extended defense is necessary for the claim that preventive war cannot be justified on the ground that those who are its targets either deserve to be killed or have consented to be killed. Because preventive war is not, by hypothesis, a response to a wrongful harm that has been inflicted but is instead based on the anticipation of future harms, it verges on incoherence to suppose that it could be justified as a form of retribution. And even if, as some have argued, combatants consent to

<sup>2</sup> Jeff McMahan, "Preventive War and the Killing of the Innocent," in Richard Sorabji and David Rodin and, eds., *The Ethics of War: Shared Problems in Different Traditions* (Aldershot: Ashgate Publishing, 2006), 169–190; and "Comment," in Michael Doyle, *Striking First: Preemption and Prevention in International Conflict*, ed. Stephen Macedo (Princeton University Press, 2008), 129–147.

become legitimate targets once war has begun, no one, to my knowledge, has argued that when people adopt the role of a soldier, they thereby consent to become legitimate targets of attack when no war is in progress and they threaten no one but are instead engaged in peacetime activities on their home bases.<sup>3</sup>

The two other forms of justification are more promising. One of these provides what I think is the best explanation of the permissibility of killing in self-defense and in defense of others. When people act in certain ways, they can lose their right not to be attacked if attacking them is instrumentally necessary to achieve a certain end and the attack is proportionate in relation to that end. When this is true, these people are, as I will say, *liable* to be attacked. Because they lack a right not to be attacked in certain ways and for certain reasons, they have no justified complaint and are not wronged if they are attacked in such a way. If, for example, a person culpably poses a serious threat to another's life, he thereby makes himself liable to be killed if that is a necessary and proportionate means of averting the threat. Although desert of harm and liability to be harmed both entail the loss of one's right not to be harmed, a liability-based justification for the infliction of harm requires that the infliction of the harm be ~~instrumentally necessary for~~ (or, in some cases, an unavoidable side effect of) the achievement of an end. By contrast, the infliction of deserved harm is, arguably, an end in itself. Traditional just-war theory can be understood as offering a liability-based justification for intentional killing in war: all combatants are liable to attack because they pose a threat to the lives or basic liberties of others and attacking them is necessary to avert the threat and proportionate in relation to the seriousness of the threat.

Traditional just-war theory offers quite a different justification for at least some of the *unintended* killing of people who are not liable to attack that inevitably occurs in war. Non-combatants are, on the traditional view, not liable to attack because they are not active participants in the war and thus do not pose a threat to their state's adversaries. Yet they are often foreseeably harmed or killed as a side effect of attacks on military targets. Just-war theory must have a justification for this, for if it is never permissible to harm or kill non-combatants in this way, the correct theory of the morality of war is contingent pacifism, not the theory of the just war. The

<sup>3</sup> Among those who have argued that combatants consent to be legitimate targets of attack in war are Thomas Hurka, "Liability and Just Cause," *Ethics and International Affairs* 21:2 (2007), 199–218, and Yitzhak Benbaji, "A Defense of the Traditional War Convention," *Ethics* 118:3 (2008), 464–495. If either of their arguments is to support traditional just-war theory, it must, it seems, assert that soldiers consent to be targets of a surprise attack that initiates a preventive war or a war of aggression.

justification that just-war theory offers is one of *necessity*. The killing of non-labile people as a side effect can be justified as necessary for the prevention of a greater evil, usually the killing of a greater number of non-labile people by one's adversary. That a necessity justification presupposes that the number of non-labile people who are prevented from being killed must exceed the number who are killed reflects an underlying moral asymmetry between killing and letting die.

A person who neither deserves nor is liable to be attacked or killed retains his right not to be attacked or killed. If the killing of such a person can nevertheless be justified on grounds of necessity, we say that his rights are overridden, or that they are justifiably infringed in the service of averting a greater evil. When the justification for attacking or killing a person is one of necessity, the person attacked or killed is *wronged*, albeit justifiably, and is entitled to compensation for what has been done to him (though that is of course difficult, though not necessarily impossible, if he is dead).

A justification of necessity is a weaker form of justification than a liability-based justification. Suppose that the lives of certain people are threatened by other people who are, because of the threat they pose, liable to be killed in defense of their potential victims. It may be permissible to kill these latter people on grounds of liability even if they significantly outnumber their potential victims. But suppose, alternatively, that the only way to prevent the original people from being killed will not kill any of the threatening people who are liable to be killed but will kill as a side effect some other people who are not liable. If this act is to be justified on grounds of necessity, those saved must, as I noted, outnumber those who would be killed. Justifications of necessity are, moreover, sensitive to the intentions with which agents act. If, for example, the only way to prevent the original people from being killed is to kill some non-labile people *as an intended means*, those saved must *substantially* outnumber those killed.

Defensive war can normally have a liability-based justification. According to traditional just-war theory, those who initiate a war pose a threat to others, thereby making themselves liable to defensive counterattack. Yet the theory of the morality of war that has most explicitly defended a liability-based account of permissible killing in war is a revisionist account according to which the basis of liability to attack in war is responsibility for a threat of wrongful harm that is serious enough to make a violent, potentially lethal attack a proportionate response. On this view, those who initiate an unjust war of aggression are responsible for a serious threat of wrongful harm and thus make themselves liable to defensive attack. This is, however, not true in the case of all defensive wars. Those who initiate a *just* war – for example,

a war of humanitarian intervention – do not pose a threat of *wrongful* harm, at least not to those whom they intentionally attack, who have made themselves liable to attack. The initiators of the just war may, of course, pose a threat of wrongful harm *as a side effect*, but if they have a necessity justification for this, that seems sufficient to exempt them from liability to defensive attack.<sup>4</sup>

But while both the traditional and revisionist accounts of the just war can thus give a liability-based justification for most defensive wars, it is less clear that either can give such a justification for preventive war. The reason for doubting whether this form of justification is available for preventive war seems obvious. Preventive war is not a response to an attack that has occurred, is occurring, or has been initiated even if no military engagement has yet occurred. It aims instead to prevent a threat from arising in the future. Yet liability derives from what a person has done or is doing; therefore there seems to be no basis for liability to preventive attack.

If there is no liability-based moral justification for preventive war, it seems that the only possible justification is a necessity justification. But preventive war involves intentional killing and for there to be a necessity justification for the intentional killing of people who are not liable, the harm that would be averted must substantially exceed that caused by the killing itself. Given that the harms that might be averted by preventive war are speculative, they must be discounted for uncertainty when weighed against the virtually certain harms that would be inflicted by preventive war. But only harms that would be near-apocalyptic in magnitude would, even after being discounted for uncertainty, *substantially* outweigh the extensive and certain harms that would be caused by a preventive war. If, therefore, preventive war requires a necessity justification, it seems likely that preventive war will in practice be seldom if ever justified.

The problem in trying to find a liability-based justification for preventive war is not that no one can be liable now to defensive action against a threat that he *will* otherwise pose later. For often those who will, unless prevented, pose a threat in the future have even now acted in ways, such as planning and preparing for a future wrongful attack, that make them liable to preventive action now. Even the mere formation of an intention to kill a person next week, or when the opportunity arises, can make a person liable to be killed. For the formation of that intention alters the objective

<sup>4</sup> See Jeff McMahan, *Killing in War* (Oxford: Clarendon Press, 2009), . 41–44; and Jeff McMahan, “Self-Defense Against Justified Threats,” in Helen Frowe and Gerald Lang, eds., *How We Fight* (Oxford University Press, [forthcoming](#)).

probabilities. It significantly increases the potential victim's objective risk of being killed. If the intended killing would be wrong and the only way to prevent it is to kill the potential murderer *now*, that person is liable to be killed and would not be wronged by being killed. Subjective and objective conditions sufficient for liability are both present: a blameworthy intention and an increase in the objective probability of a person's being wrongly killed. One might argue that killing the potential murderer in advance of his actually acting on the intention he has formed would be to fail to respect him as an autonomous person by denying him the opportunity to change his mind. But it is highly doubtful that an innocent person ought to bear a significantly increased risk of being murdered in order to enable a person who wrongfully intends to commit a murder to have the chance to abandon his intention. By forming that intention, the potential murderer has wrongfully created a situation in which, by hypothesis, either he must be killed or his potential victim must be exposed to a high risk of being murdered. Justice requires that the one who is morally responsible for this situation bear the cost.

The same is true even when a potential murderer has not formed an intention to kill but is seriously considering whether to murder someone. Even to deliberate about whether to commit a murder is blameworthy and increases the objective risk that the potential victim will be murdered. Suppose that before the potential murderer conceived the idea, the objective probability that the potential victim would be murdered was near zero. While the potential murderer deliberates, the objective probability rises to 10 percent. If the *only* opportunity for intervention is *now*, why should a wholly innocent person have to bear a 10 percent risk of being murdered in order that the person who has wrongfully created that risk should be spared?

In practice, of course, these reflections are idle, since an intention alone without any external manifestation cannot be detected and even if somehow the intention becomes known, there are usually ways, which I excluded in the example by mere stipulation, of preventing a future murder other than killing the potential murderer well in advance. The point is only that it is in principle possible for a person to be liable to attack now to prevent him from doing that he would otherwise do only much later.

This may not, however, provide a secure foundation for a liability-based justification for preventive war. For those who are culpably responsible in advance for an increased risk of future aggression are the political leaders who conceive of, deliberate about, intend, and plan and prepare for such a war. For a variety of reasons, such people generally are not and cannot be

the targets of preventive war. The targets must instead be those who would later carry out the orders from the political leaders to fight a war of aggression – namely, soldiers who at the time are engaged in peacetime activities and may be entirely unaware of the intentions of their leaders.

Many proponents of traditional just-war theory can be expected to resist the suggestion that the theory can offer only a necessity justification for preventive war and thus effectively rules out the prevention of future aggression as a just cause for war. They can point out that the traditional theory makes the principles of *jus in bello* wholly independent of the principles of *jus ad bellum*, so that what it is permissible for combatants to do in war is independent of whether the war itself is just or unjust. But if this is so, it seems that the theory implies that unjust combatants do no wrong merely by fighting, even in an unjust war of aggression, provided they confine their attacks to military targets, do not mistreat prisoners of war, ensure that their attacks are proportionate to their military objectives, and do not exceed the degree of force necessary to achieve those objectives. This is true, these theorists might point out, even when unjust combatants initiate an unjust war of aggression with a surprise attack against unmobilized soldiers on their bases. For a military base is a military target. If unjust combatants do no wrong in attacking a military base during the course of a war, they also do no wrong if they initiate a war with such an attack. No traditional just-war theorist has, to my knowledge, held that *individual combatants* are guilty of wrongdoing if they initiate a war with a surprise attack on unmobilized soldiers, though all of course claim that the political leaders who order a surprise attack to initiate an unjust war of aggression are guilty of a heinous moral wrong. No one, in other words, suggests that a surprise attack that initiates a war is the sole exception to the independence of *jus in bello* from *jus ad bellum* – the sole instance in which an attack that harms only soldiers who have been neither injured nor captured constitutes a moral wrong or, morally speaking, a war crime.

It therefore seems that traditional just-war theory should be able to find a justification for at least some instances of preventive war that is not a necessity justification. For if it can be permissible for combatants to attack unmobilized soldiers in initiating an unjust war of aggression, surely it can also be permissible for them to attack unmobilized soldiers as the first act of a preventive war that may actually prevent unjust aggression in the future. Yet the assumption that it is open to traditional just-war theorists to advance this argument is an indication not so much of the range of the traditional theory's resources as it is of the theory's possible incoherence. For it is doubtful that the theory can consistently concede that it is morally

permissible for combatants to initiate an unjust war – or indeed a just war – by conducting a surprise attack on unmobilized soldiers.

Recall that I noted earlier that traditional just-war theory asserts that all combatants are legitimate targets in war because they are liable to defensive attack, and they are liable because they actively pose a threat to others. This idea is presupposed by the pervasive use of “innocent” to refer to those who are *not* legitimate targets, together with the explicit identification, in many documents in the just-war tradition written after the early “classical” period, of the *innocent* with those who are *unthreatening* – an identification licensed by the etymology of the term. But if liability is a function of posing a threat and unmobilized soldiers on their home bases pose no threat, they cannot be liable to attack even if they are “military” and thus constitute a military target.<sup>5</sup>

Traditional just-war theory therefore faces a dilemma. Either unmobilized soldiers are liable to attack and thus are legitimate targets or they are not. If combatants are legitimate targets in war only because they pose a threat to others, and if unmobilized soldiers in peacetime do not pose a threat to others and are thus not combatants in the relevant sense, then unmobilized soldiers are not legitimate targets. There is much to be said for this claim. Unmobilized soldiers in peacetime would seem to be *hors de combat* in the relevant sense, for there is no combat in which they are participants. But in that case, individual combatants act wrongly if they participate in the initiation of either an aggressive war or, in the absence of a necessity justification, a preventive war that requires an attack on unmobilized soldiers. While it may seem plausible to suppose that it is wrong to participate in an attack that initiates an unjust war of aggression, this supposition is incompatible with the traditional theory’s claim that it is not impermissible to fight in an unjust war provided one fights in accordance with the rules. If just-war theory were to concede that it can be wrong to participate in an attack against unmobilized soldiers that initiates an unjust war, that would, as I noted above, be tantamount to its accepting that there is one exception to the independence of *jus in bello* from *jus ad bellum*.

If, by contrast, unmobilized soldiers engaged in unthreatening activities on their bases in peacetime are nevertheless legitimate targets of attack, just-war theory cannot continue to claim that the sole basis of liability to

<sup>5</sup> I pressed this challenge some years ago in McMahan, “Innocence, Self-Defense, and Killing in War,” *Journal of Political Philosophy* 2:3 (1994), 193–221; 196–198, but to my knowledge it has never been discussed in the just-war literature.



attack is posing a threat to others. “Innocent” cannot mean “unthreatening,” and the categories “innocent” and “noncombatant” cannot be assumed to coincide. If people can be liable to attack in war for reasons other than that they pose a threat, it may then be possible to justify preventive war without appealing to a necessity justification. But if traditional just-war theorists wish to adopt this option while retaining the essential elements of their view, they must identify a criterion of liability to attack that implies that while unmobilized soldiers and combatants in war are both legitimate targets, noncombatants are not.

Since just-war theory has tended to assert that combatants act permissibly if they confine their attacks to military targets, the obvious candidate for the criterion of liability to attack in war is membership in the military. But that has the implausible implication that if one were to go onto a military base in peacetime and attack someone in uniform, one would not have wronged that person, who would have no justified complaint about what one had done. Perhaps the appropriate response to this is to claim that the liability of military personnel is not liability to attack *by anyone* but liability to attack only by enemy combatants. But this is still not right, since it implies that if a soldier of one state, acting on orders, were to go onto a military base in another state during peacetime and attack someone in uniform, the person attacked would not be wronged. To avoid this counter-example, the alternative criterion of liability needs to be qualified so that it applies only in wartime, or while war is in progress. This yields an understanding of liability that should satisfy the demands of traditional just-war theory – namely, that a person who is a member of the military is liable to attack by enemy combatants acting under orders during wartime.

There are, however, various objections to this account of liability to attack. They can be divided into two broad challenges: (1) that the account implies that much that is morally significant depends on whether the act occurs *during wartime*, and (2) that the account makes liability a matter of membership rather than action. I will consider these objections in turn.

If membership in the military makes a person liable to attack only when his state is *at war*, then in order to know whether a person is liable to attack at a particular time, one must have a criterion for determining whether a conflict is a *war* or some lesser form of conflict, as well as a criterion for determining precisely when wars begin and when they end. It may seem obvious that if unmobilized soldiers on their base are killed in a surprise attack by combatants of a state with which their state has not been at war, they cannot have been liable to attack. For they were attacked when their state was not at war – that is, at a time when no war was in progress.

Yet the just-war theorist who wishes to argue that preventive war can be justified without an appeal to necessity might claim that the surprise attack itself constitutes the first act in a war, so that the unmobilized soldiers were in fact killed *in war*, though admittedly very early in the war, and so were liable to attack.

There are at least two strong objections to this suggestion. First, it presupposes that the political leaders on the side that initiates a war have remarkable powers of moral alchemy. By declaring that the surprise attack is an act of *war*, they seem to be able to strip unmobilized soldiers of their right against enemy soldiers not to attack them, making them liable to attack without their having done anything to change their moral status. Suppose, for example, that the leaders and their agents who launched the attack were spectacularly obtuse and intended it not as an act of war but only as a test of their missiles. In that case, the soldiers killed might not have been liable, since they would have been killed as a side effect of a missile test rather than in a war. Of course, the leaders of the state whose soldiers were the victims of the attack might not care what the other leaders' intentions were. They might go to war in retaliation even if they accept that the other leaders simply acted callously without intending to start a war. Or they might not. This is the second objection: that whether or not the surprise attack is the opening act of a war may depend on how the state that is the victim of the attack decides to respond. Suppose that the state that has been attacked refrains from retaliating and instead pursues diplomatic activity that is later rightly credited with having averted a war. In that case the surprise attack was not an act in a war but an attack that might have but did not precipitate a war, so that the victims of the attack were not liable according to the criterion of liability that seemed so promising for traditional just-war theory. That criterion thus implies that the victims of a surprise attack also have powers of moral alchemy: they can determine, by their response to an attack, whether those killed were liable to be killed or whether, instead, the killings were murders.

The second broad objection to the claim that members of the military are liable to attack only by enemy combatants and only in war is that it is ad hoc and devoid of moral significance. Although it is a mistake to suppose that posing a threat to others is a sufficient condition of liability to attack, one can at least see how that criterion has *prima facie* moral significance. But what is the significance of membership in the military in a time of war, given that many members of the military pose no threat in war and bear no responsibility for the acts of their fellow members who do? This criterion makes liability a matter of membership rather than of action.

Or perhaps the claim that members of the military are legitimate targets for opposing combatants in war is not a claim about liability at all. It might be more plausibly regarded as a claim about the instrumental value of identifying certain sharply defined categories (“soldier” and “civilian”) that are widely *perceived* to have moral significance, and assigning the members of the different categories different moral, conventional, or legal statuses. This is in fact one way in which certain prominent just-war theorists have sought to defend the claim that all soldiers in war are liable to attack (though they sometimes run this instrumental justification in tandem with the claim that soldiers are legitimate targets because they threaten people they have been trained to harm).<sup>6</sup> This is a form of justification for killing people – that it is morally permissible to kill certain people if it would be better for people generally if they were to agree to act on the assumption that it is permissible to kill these people – that I will not consider here, other than by making two points: first, that it is a form of justification I do not accept, in part because it makes our moral status contingent on extrinsic circumstances, and, second, because even if it were true, I doubt that it would support the idea that all soldiers, but no civilians, are legitimate targets in war. But these are topics too large to be addressed here.

The conclusion I draw from this discussion of the traditional theory of the just war is that it seems unlikely that that theory has the resources to support a liability-based justification for preventive war, at least when preventive war would involve attacks on unmobilized soldiers, who have not been involved in the planning and preparation for the aggression that the war would be intended to prevent. Traditional just-war theory must either condemn preventive war or appeal to one or the other of two different forms of justification: a necessity justification or a justification that claims that it is instrumentally valuable to treat unmobilized soldiers as legitimate targets.

Earlier I noted that there is a revisionist account of the just war according to which the criterion of liability to defensive attack is responsibility for a threat of wrongful harm that is sufficiently serious to make killing a proportionate response. This account, I suggested, can give a liability-based justification for at least some instances of preventive attack against those who intend, plan, and prepare for an unjust attack. But can it supply a liability-based justification for preventive attacks against

<sup>6</sup> For a recent example, see Avishai Margalit and Michael Walzer, “Israel: Civilians and Combatants: An Exchange,” *The New York Review of Books* 56:8, May 14, 2009, 21–22. For an able defense of a related view, see Benbaji, “A Defense of the Traditional War Convention.”

unmobilized soldiers, particularly if they are non-culpably ignorant of their government's plans for aggression? Because this revisionist account makes liability a matter of action rather than membership, it implies that some members of the military who make no contribution to an unjust war may not be liable to attack even when the war is well underway. It may therefore seem unlikely that it could recognize unmobilized soldiers as liable to preventive attack. Yet in the earlier work to which I referred at the beginning of this essay, I argued that unmobilized soldiers can be liable to preventive attack if two conditions are satisfied: (1) that they chose to become members of the military, even if they did so under a certain degree of compulsion, and (2) that there is a substantial probability that they will obey an order to fight if it is given.

On its own, the fact that there is a high risk that a person will pose a threat of wrongful harm in the future – supposing this fact were knowable – seems insufficient for liability to preventive attack. But if, in addition to this, the person has made a choice that had as a foreseeable risk that he would later pose a threat of wrongful harm, he may then be liable to preventive attack *provided* that the likelihood of his becoming a threat is traceable to his having risked becoming one. His choice to risk becoming a threat may not be a basis of liability if the reason he is likely to become a threat is independent of that choice. But in actual cases, the reason why unmobilized soldiers contribute to the risk that their leaders' plotting imposes on others derives from their being the ones who will later implement their leaders' plans. That they might later be in this position, even without knowing it at the time, was foreseeable when they chose to become members of the military.

One reason why there is normally a substantial probability that soldiers will fight if ordered to do so is that in most cases, they precommit their wills to obedience when they join the military. They know that this is expected of them and in most cases there is even a formal precommitment at induction in the form of an oath of obedience, which tends to strengthen their sense that they have a moral duty of obedience. The precommitment of the will is then reinforced by the knowledge that virtually all of those who join have that there are severe penalties for disobedience. For these and other reasons, the act of joining the military tends to create a high probability that the soldier will fight if ordered to do so, even if the war in which he is ordered to fight is unjust, and even if he recognizes this.

When it is true that a soldier is strongly predisposed to fight if ordered to do so, and when his leaders are in fact planning to order him to fight in an unjust war, his presence in the military increases the objective risk faced

by the potential victims of this potential unjust war. If he has chosen to join when these conditions were among the foreseeable possible outcomes of his joining, that seems to provide a basis of liability to preventive attack, as a means of reducing the risk to which his action has contributed.

An initial objection to this explanation of how unmobilized soldiers can be liable to preventive attack is that many members of the military do not become members by choice, so that in their case the first of the two conditions of liability is not met. But in fact almost all members of the military are members by choice, even if the choice is made under duress. It is, of course, possible that a person's government has the authority to induct him into the army against his will, assign him a rank, and lock him in a military prison if he resists. In that case he could be a member without any choice on his part. Anyone who has become a member of the military in this way, without any element of choice on his part, cannot be liable to preventive attack on the basis of his membership. But soldiers rarely become members of the military in this way. Usually they enter the military by choice, even if their choice is made under considerable pressure. People who make choices under duress may be exempt from *culpability* for their choice, but that is not the same as being exempt from *responsibility*. Provided they could have chosen differently, they may be responsible for the consequences of their choice even if, in the same circumstances, most people of reasonable firmness would not have chosen differently. And responsibility can be a basis of liability, even in the absence of culpability. This is most obvious in the case of liability to pay compensation for having caused harm. A person who chooses under duress to harm an innocent person, and lacks a justification of necessity for the harm caused, may clearly be liable to compensate his victim, even if the duress made his choice non-culpable.

While liability to compensate a victim for a harm caused does not entail liability to defensive action *ex ante*, there are examples that suggest that responsibility can also be a basis of liability to defensive action even in the absence of culpability. Here is one I have used in previous work to illustrate this point.

*The Conscientious Driver:* A person who keeps her car well maintained and drives carefully and alertly decides to drive to the cinema. On the way, a freak event that she could not have anticipated occurs that causes her car to veer out of control in the direction of a pedestrian.<sup>7</sup>

<sup>7</sup> McMahan, *Killing in War*, 165.

Suppose the pedestrian can defend himself, but only by killing the driver. Given that it is unavoidable that one of them will be killed, it seems that the driver's having chosen to engage in an activity that is known to involve a small risk of killing an innocent person makes it just for the pedestrian to kill her, provided there are no other considerations that weigh heavily against the pedestrian's acting in self-defense. The driver's earlier voluntary choice to impose the risk on others provides the basis of her liability to defensive action.

Not everyone, of course, agrees with this judgment. And there are several morally significant differences between the driver in this case and unmobilized soldiers whose government is secretly planning an unjust war of aggression. One such difference suggests that the grounds for attributing liability to preventive attack to unmobilized soldiers are actually stronger than those for attributing liability to defensive attack to the conscientious driver. This is that what the soldiers risk doing is more seriously wrong than what the driver risks doing. Whereas the driver risks killing innocent people accidentally, the unmobilized soldiers risk being in a situation in which they will *intentionally* kill people – soldiers who seek to defend their country against unjust aggression – who are innocent in the sense identified as relevant by the revisionist account of the just war. They risk being in a position in which they will kill such people either because they will mistakenly believe that the victims are not innocent or because they will lack the strength of will to refrain from killing them despite being aware that they are innocent.

There are, however, other differences between the cases that suggest that the grounds for attributing liability to the conscientious driver are stronger. One is that while the conscientious driver freely chose to drive for reasons of self-interest, many unmobilized soldiers may have joined the military only under duress. In those cases, the duress makes a difference to the degree of the soldiers' responsibility for their contribution to the threat of unjust war, thereby diminishing the degree of ~~their liability to preventive attack.~~

But not all unmobilized soldiers join the military under duress. Many enlist voluntarily, for a variety of reasons. Some are culpable for enlisting – for example, if their government has a history of flagrantly and egregiously immoral action which makes it reasonable to expect that it is likely to use its military to wage unjust war. In such cases, it is not difficult to find grounds for attributing liability, even to preventive attack, to those who have willingly converted themselves into instruments of violence in the service of notorious wrongdoers.

Others who enlist voluntarily do so for admirable and even noble reasons. They may join, for example, knowing that doing so puts them at considerable personal risk, but in the *reasonable* expectation that they will be ordered to fight only in wars that are just. They may, in other words, expose themselves to substantial risk in order to help defend innocent people against wrongful attack. It is hard to see how they could thereby make themselves liable to attack. As I noted earlier, it seems that action done with moral justification – that is, action that is permissible and that there is a positive moral reason to do – is not a basis of liability to defensive harm. If so, unmobilized soldiers who were morally justified in joining the military cannot be liable to preventive attack simply by virtue of having joined. In this respect they contrast with the conscientious driver, whose driving is morally permissible but not morally justified, in that there is no positive moral reason for her to drive to the cinema. This may explain why we might think that she is liable to defensive attack while unmobilized soldiers who had a positive moral reason to join the military are not. Mere permissibility, without positive moral justification, seems insufficient to exempt a person from liability.

Yet given that their government is in fact planning and preparing for an unjust war of aggression, it seems that their joining the military, or remaining in it, was not justified in what Parfit calls the “fact-relative” sense – that is, justified in relation to the facts. At best their action was justified in the “evidence-relative” sense – that is, in relation to the evidence available to them.<sup>8</sup> While they may have acted in good faith on the basis of beliefs they were epistemically justified in having, some of the beliefs they had that were relevant to the justifiability of their action were false, so that their action turned out to be unjustified in the fact-relative sense. It is only justification in the fact-relative sense that excludes liability. While evidence-relative justification is sufficient to exempt unmobilized soldiers from culpability for having joined, it is insufficient to exempt them from liability to preventive attack.

Some philosophers believe that mere evidence-relative justification does exempt an agent from liability to defensive harm. But many of these same philosophers reject my claim that fact-relative justification excludes such liability. They argue, for example, that a pilot who has a fact-relative lesser evil justification for dropping a bomb on a military target when that will foreseeably cause unavoidable but proportionate harm to innocent civilians

<sup>8</sup> Derek Parfit, *On What Matters*, vol. 1 (Oxford University Press, 2011), chapter 7, section 21.

can be liable to defensive attack.<sup>9</sup> While I think this particular claim is false, I concede that there are cases in which justification does not exclude all forms of liability. Suppose, for example, that through no fault of his own, a person has gone into a diabetic coma and will soon die unless he is administered an injection of insulin. One has no insulin oneself but one can steal the necessary amount from someone else who has an ample supply. It seems clear that one would be morally justified in stealing the necessary amount of insulin but that one would then be liable to compensate the original owner for having stolen his property. It might be fairer, all things considered, if the owner of the insulin were to waive his right to compensation, or if the diabetic who benefited from one's intervention were to compensate him instead. But if neither of these people volunteers to bear the cost, it seems that one is liable to compensate the person from whom one has stolen, despite the fact that one acted with moral justification.

It seems, therefore, that my earlier claim that justification excludes liability was, at a minimum, overstated. Justification does not always exempt one from liability to compensate innocent people harmed by one's action. Yet even if a moral justification for acting is compatible with liability to compensate innocent victims of the action, the justification may still exclude liability to *defensive* harm. It seems to be true, for example, that in justifiably stealing the insulin, one would not be liable to defensive action, even though one would be liable to compensate the owner. So it may be that justification always excludes liability to defensive action, even if it does not always exclude liability to compensate innocent victims. (The explanation of why one would not be liable to defensive action in the insulin case might be that the circumstances had deprived the owner of his liberty-right to keep the insulin, so that one would be forcing him to do what he ought to do. But he might still retain a claim-right to keep the insulin – an instance of a right to do wrong. The theft is justified because that claim-right is overridden, yet the overriding of that right may be what gives the owner a claim to compensation.)

Even if, as the insulin case suggests, fact-relative justification (either a liability justification or a lesser evil justification) exempts an agent from liability to defensive attack, it is, I suggested, considerably less plausible to suppose that mere evidence-relative justification does so. But unmobilized soldiers whose government is secretly preparing for an unjust war of

<sup>9</sup> See, for example, Uwe Steinhoff, "Jeff McMahan on the Moral Inequality of Combatants," *Journal of Political Philosophy* 16:2 (2008), 220–226; and Adam Hosein, "Are Justified Aggressors a Threat to the Rights Theory of Self-Defense?" forthcoming in Frowe and Lang, *How We Fight*.



aggression seem to have had only an evidence-relative justification for joining the military, no matter how admirable their motives may have been. Their having joined has placed them now in a position in which there is a very high probability that they will participate in an unjust war of aggression unless they are forcibly prevented from doing so. Since this was a foreseeable possible result of their joining – for everyone knows that soldiers often fight in unjust wars and that very few ever refuse an order to do so – their having joined seems to constitute a basis of liability to preventive attack.

Some people's intuitions will rebel at this suggestion. Suppose that at the time they joined the military, a group of unmobilized soldiers were epistemically justified in believing that their government would fight only in wars that would be just. Yet now, contrary to reasonable expectation and unbeknown to them, their government will soon order them to fight in an unjust war of aggression. In all that they have actually chosen and done, they may be morally indistinguishable from unmobilized soldiers on the other side who may soon be ordered to conduct an attack against them to prevent them from later engaging in unjust aggression. It therefore seems as if the account of liability I have defended implies that they have become liable to attack by bad luck alone. That seems an arbitrary basis for the attribution of liability.

Indeed, it *is* an insufficient basis for the attribution of liability, *on its own*. But recall that my claim is that there are *two* conditions that must be met for unmobilized soldiers to be liable to preventive attack. One is they chose to join the military when there was a foreseeable risk that that would result in their fighting in an unjust war; the other is that they would in fact obey the order to fight in an unjust war that they will receive unless they are preventively attacked. While their prior choice to join the military may be a necessary condition of their being liable to preventive attack (and thus distinguishes them from civilians who will soon receive notification of conscription and thus will also pose a threat of unjust harm in the future), it is not sufficient. There is a further condition – what they would do if ordered – that is also necessary for liability. For unlike the conscientious driver, the unmobilized soldiers would have to make a further choice in the future before they would actually pose an unjust threat. They would have to choose to fight when ordered to do so.

They have not, *ex hypothesi*, made that choice; not yet. And those who would not make it – that is, those who would not be ordered to fight or would refuse to obey the order – are not now liable to preventive attack. For they fail to satisfy the second of the two conditions of liability. This

second condition is essential, as it distinguishes those who would later pose a threat from those who would not. And the first condition is essential as well, as it distinguishes soldiers from high school students in a country with universal conscription, of whom it may be true that they later will pose a threat but who as yet have made no choice that explains why they may become a threat.

Even though unmobilized soldiers have not as yet chosen to fight, it is statistically certain in advance that virtually all of them will fight if they receive the order. This is what virtually all soldiers have always done. And we understand the reasons why they have done so and will continue to do so unless the relevant conditions change. They obey when ordered to fight because they have precommitted their wills to obedience, because they have been conditioned to obey, and because they believe, however mistakenly, that their war is just. *Or*, even if they recognize that it is unjust, they will nevertheless obey because they would be harshly punished and ostracized if they do not, because patriotism and a sense of professional and contractual obligation impels them to obey, and so on. I knew a full colonel in the US army who told me, when the war in Iraq had been in progress for a couple of years, that he believed with great conviction that the war was unjust and that he despised the Bush administration for having started it, but that he would fight to the best of his ability if he were deployed to Iraq. His view was unusual among the army officers I knew at the time only in being volunteered without prompting or hesitation.

It seems, therefore, that when unmobilized soldiers will otherwise receive an order to fight in an unjust war of aggression, most of them are liable to attack if that is the only way, or even just the best way, to prevent them from engaging in an unjust attack. This claim is not, moreover, restricted to soldiers who will be ordered to engage in external aggression. It applies equally to soldiers who will be ordered to engage in domestic repression. As I am making final revisions to this chapter, soldiers in the Syrian army are attacking and killing protesters against the government in various cities throughout that country. Suppose that in the evening one were to learn that President Assad had just sent an order that a particular unit of the army would receive the next morning to attack certain neighborhoods in Hama. And suppose that although one would be unable to attack the members of that unit once they had begun their assault, one could attack them preventively during the night. The account I have defended implies that they would be liable to that attack.

In many cases of preventive attack, however, there may be a proportion of those attacked who are not liable because they would not later pose a

threat of unjust attack. Some would have left the military by the time the order was given; others would already be assigned to tasks unrelated to the war that would occupy them for its duration; some would be ill or otherwise incapacitated; and some might refuse on moral grounds to fight in an unjust war of aggression. These soldiers would be innocent in the relevant sense, yet a preventive attack on them would appear to involve killing them intentionally. Could that be morally justified, and if so on what grounds?

Although this may seem to be an issue of discrimination, it is actually an issue of proportionality. The requirement of discrimination is a moral constraint on the intentional killing of innocent people. In intentionally attacking a group of unmobilized soldiers, some of whom one knows to be innocent in the relevant sense, one might seem to be intentionally killing at least some people one knows to be innocent, so that one would be in violation of the requirement of discrimination. But this is a mischaracterization of what one would be doing – or at least of what one *might* be doing.

Traditional just-war theorists concede that it can be permissible to kill innocent civilians as a foreseen side effect of intentionally attacking a military target, provided that the killings are not disproportionate in relation to the importance of destroying the target. Suppose, for example, that a pilot fighting in a just war sees a large concentration of enemy ground forces in the center of a large open area. There are hundreds of them, all en route to his nation's capital to participate in a vicious siege and all currently vulnerable to attack. He also sees, however, that they are holding a few civilian hostages; but he correctly concludes that bombing the unjust combatants is sufficiently important to make the unintended killing of the innocent hostages a proportionate side effect. It may seem, of course, that bombing this group that includes both a great many people who are liable to be killed and a few who are not is relevantly different from bombing a group of unmobilized soldiers that also includes some who are liable and others who are not. For the pilot in the former case can identify the individual civilian hostages and exclude them from among those he intends to kill, whereas a pilot attacking the unmobilized soldiers cannot tell which of them are not liable to attack and so must intend to kill them all, including those who are in fact innocent, or not liable.

But suppose that the first pilot, having surveyed the concentration of enemy forces, now circles back to drop his bombs, only to discover that the civilian hostages have been hastily dressed in combatants' uniforms in the hope that the pilot will believe that he can no longer conduct a discriminate attack, since he will now have to intend to kill all of the people in

uniform, knowing that a few of them are innocent. But the enemy forces have underestimated the pilot's philosophical intelligence. He knows that his intentions have not changed. He can still drop his bombs without intending to kill innocent people. It is only that he is now unable to identify those whose deaths he does not intend. It does not matter whether his inability to identify them results from their being hidden inside a building or their being hidden inside a uniform.<sup>10</sup>

For both pilots, the only relevant issue is one of proportionality. Each recognizes that he has no reason to kill those who are not liable, for neither the civilian hostages nor the unmobilized soldiers who would not fight if their country were later to initiate an unjust war of aggression constitute a threat; hence killing them would serve no purpose. What each pilot needs to know to determine whether it would be permissible to attack is just roughly how many innocent people he would kill, and perhaps the rough proportion among the people he would kill who would be liable to be killed. This is a question about whether his attack would be proportionate, not about whether it would involve the intentional killing of the innocent.

There is a further issue that is relevant to the moral status of unmobilized soldiers who, for whatever reason, would not later fight in the unjust war for which their government is currently preparing. This is that in having chosen to join the military, they are to some degree responsible for having engendered a reasonable expectation among their potential adversaries that they will fight if ordered to do so. They are therefore responsible in some measure for *appearing* to pose a threat of unjust harm when those who are the potential victims of their government's planned war of aggression discover the plan.

There are at least three ways in which this fact might be morally significant. One is that this fact is sufficient to make these unmobilized soldiers liable to preventive attack along with the others who *would* fight. Consider a case involving only two individuals that is analogous except that it involves apparent defense rather than apparent prevention. Suppose the friends of a person who is known to relish the terrified excitement he experiences while watching trashy horror movies decide to organize a treat for him. They arrange to pay a particularly large and powerfully built actor to dress up in the usual regalia of the Hollywood killer (leather mask, etc.) and to pretend, as realistically as possible, to be intent on sadistically killing their friend. The plan is that just as this terrifying figure appears to be

<sup>10</sup> The arguments in these paragraphs are drawn from McMahan, "Innocence, Self-Defense, and Killing in War," 215–221, and *Killing in War*, 229–230.

preparing to strike the fatal blow, the friends will emerge from their hiding places to reveal that it was all just a special treat. But the best-laid schemes, this one among them, gang aft agley. When the pretended killer appears, the horror movie fan, reasonably believing he is about to be murdered, fires the derringer he has unexpectedly concealed in his pocket, killing the actor before he has even had the chance to utter a fiendish taunt. The question is whether the actor had actually made himself liable to be killed by deliberately creating conditions that were indistinguishable, from the point of view of the fan, from a situation in which he was actually about to be murdered. One might argue in the following way that he had. Suppose that an armed and fully informed third party improbably arrives on the scene just as the fan is about to kill the actor. This person knows both that the actor poses no threat and that the fan has a gun and is about to fire it in what he believes to be justified self-defense. In these circumstances, a killing is unavoidable. The third party must choose between killing the fan in defense of the actor and allowing the fan to kill the actor. One might initially think that because the fan has chosen to kill someone but has made a mistake about the justification for that choice, he bears responsibility for the situation in which a killing is inevitable and on this basis is liable to be the one who is killed. But on reflection it seems that the actor's responsibility for the situation is greater, for he has knowingly produced the fan's mistaken belief. Given that it is unavoidable that one of them will be killed, the third party ought to act to ensure that it is the actor rather than the fan, on the ground that the actor bears greater responsibility for its being inevitable that one of them will be killed. This has the appearance of a liability-based justification for non-intervention by the third party.

It is hard to believe, however, that the actor could make himself morally liable to be *killed* simply by trying to frighten someone, however reckless or imprudent the attempt may have been, given that he does not in fact pose a threat. To the extent that it is credible to suppose that he could be liable, it is because his action was culpable.<sup>11</sup> But unmobilized soldiers who were justified in the evidence-relative sense in joining the military and would not later fight are neither culpable nor threatening. It is thus even harder to believe that they could have made themselves liable to be preventively

<sup>11</sup> My thinking here is indebted to Kai Draper's criticism of some of my earlier work in his paper, "Defense," *Philosophical Studies* 145:1 (2009), 69–88, esp. 74 and 86–87. In McMahan "Self-Defense and Culpability," *Law and Philosophy* 24:6 (2005), 751–774, I offer intuitive grounds for the claim that culpability alone can be a basis of liability to be killed in defense of another, even if one poses no threat oneself. My argument there may support the view that the actor is liable to be killed.

killed simply by having knowingly made it reasonable for their potential adversaries to believe that they would fight if ordered to do so.

But even if their responsibility for appearing to pose a threat of unjust harm is insufficient to make them liable to preventive attack, it might be relevant in a second way. Assuming that they are not liable to attack, they would be wronged by being attacked. But because they have chosen to be in the military, thereby knowingly making it reasonable for others to suppose that they will fight if ordered to do so, they lack a right to kill in self-defense if others attack them on the basis of the reasonable and epistemically fully justified belief that they will soon receive an order to fight in an unjust war of aggression and will respond with obedience. Because they are themselves responsible for their attackers' mistaken belief, they must accept the unavoidable cost of having made that belief reasonable, rather than imposing the cost on the same innocent people on whom they imposed the belief.

The third possible implication of the unmobilized soldiers' responsibility for appearing to pose a threat is that it diminishes the extent to which they are wronged by being preventively attacked when the reason they are attacked is precisely that they present that appearance, thereby giving their attackers an evidence-relative justification for attacking them. Even if they are not liable to attack, they may have little ground for complaint about being attacked if they have knowingly chosen to make it reasonable for their attackers to believe that they may justifiably be attacked. If this is so, it seems it must be at least relevant to the weight their deaths should have in determining whether a preventive attack would be proportionate.

Having identified three ways in which the responsibility of unmobilized and unthreatening soldiers for appearing to pose a threat might be morally significant, I should concede that whether this responsibility actually makes them liable, whether it actually deprives them of the right of self-defense, and whether and to what extent it actually diminishes the extent to which they are wronged by being preventively attacked may depend on the *degree* of their responsibility for the appearance they present, which in turn depends on the degree to which their joining and remaining in the military have been voluntary. If they entered and have remained in the military only under great duress, that mitigates their responsibility for appearing to pose a threat when in fact they do not and will not.

Even when their entry into the military was fully voluntary, the fact that soldiers who will not fight are nevertheless responsible for appearing as if they will is at best a weak basis of liability to preventive attack. The principal conditions of liability, according to the revisionist account of

the just war, are the following. First, soldiers have chosen to be in the military, thereby granting their government the authority to order them to fight and usually precommitting their wills to fight if ordered to do so. Second, they would in fact fight if ordered to do so – a condition that it is reasonable to expect to be met in most cases, for the various reasons I gave earlier. Third, their government will soon, unless prevented, order them to fight in an unjust war of aggression. When these three conditions obtain, unmobilized soldiers can be liable to preventive attack. There may always be some who fail to satisfy the first or second conditions – for example, those who have been forcibly dragooned into the military, or who will no longer be in the military when the order to fight is given, or who will receive that order but conscientiously refuse to obey it. Unless these soldiers can be liable to preventive attack merely on the ground that they are responsible for appearing to pose a threat, there can be no liability-based justification for killing them. It does not follow, however, that preventive attack cannot be justified. For the foreseen killing of unmobilized soldiers who are not liable to attack may be justifiable in the same way that the foreseen killing of innocent civilians as a side effect of military action can sometimes be justified in the course of just warfare. In other words, the killing of unmobilized soldiers who are not liable may be justifiable provided that it is not intended as a means to victory, is unavoidable in the circumstances, and is proportionate in relation to the moral importance of the military goal. This is a necessity justification, which supplements the liability-based justification for killing those soldiers who satisfy the three conditions of liability. According to the revisionist account of the just war, then, the primary justification for preventive attack – the justification for the killings that are intended – can be liability-based, though this justification must often be supplemented by a necessity justification for the unavoidable killing as a side effect of people, including some soldiers, who are not liable to attack. As in the case of most justified military action in modern war, considerations of liability justify the harms that are intentionally inflicted, while considerations of necessity or lesser evil justify those that are unintended.

Of the three conditions of liability, two require knowledge of one's adversary that it is difficult to have – namely, that the government is planning an unjust war and will at some point order its soldiers to fight, and that those soldiers will obey the order when they receive it. But the relevant knowledge is not impossible to obtain. Knowledge of a government's secret plans may be obtained by espionage and knowledge of what soldiers will do is given by history and psychology. Conditions would be

quite different, however, if there were no reasonable expectation that virtually all soldiers will obey an order to fight in an unjust war of aggression. If people's understanding of the morality of war were to shift away from the traditional just-war doctrine, which holds that soldiers do no wrong in fighting in an unjust war, provided they obey the rules governing the conduct of war, so that people came to believe instead that it is seriously wrong for soldiers to kill people in pursuit of unjust aims, then soldiers might become reluctant to fight in wars they believe to be unjust. If the consensus about the morality of war were to shift in this way, and if this change were to lead to the adoption by civilized peoples of more generous legal provisions for conscientious refusal to fight, it might become much more difficult for the revisionist account of the just war to provide a liability-based justification for preventive war.