

Hobbesian Defenses of Orthodox Just War Theory

1 Orthodox Just War Theory

Most of us accept that all persons have a right not to be killed unless by their action they have forfeited it, and that there is thus a strong constraint against killing people unless they have forfeited or waived that right. According to the currently dominant understanding of the just war, civilians retain the protection of this right in conditions of war but combatants do not. On one view, combatants forfeit the right by posing a threat to others; on another, they waive it when they accept combatant status, which requires that they identify themselves visually and in other ways as legitimate targets. Yet people who fight in a just war (“just combatants”) and fight only by permissible means, are simply defending themselves and other innocent people against a wrongful attack or some other serious wrong. There seems to be no reason to suppose that they thereby either forfeit their right not to be killed or grant their enemies permission to try to kill them.

It seems, therefore, that the permission that those who fight without a just cause (“unjust combatants”) have to kill just combatants is a *legal* permission only, not a moral permission. If so, the law of war diverges quite radically on this issue from the morality of war. Although just combatants retain their moral right not to be killed, and although their right is seldom overridden, it is nonetheless best, for a variety of contingent and largely pragmatic reasons, not to hold unjust combatants legally liable for killing them. The moral right of just combatants not to be killed is not and, at least in current conditions, should not be protected by a legal right. The killing of just combatants in war should not be treated as murder.

This is, however, not the common view of the permissibility of killing just combatants in war. Most people, including most contemporary moral theorists writing in the just war tradition, believe that the morality of war and the law of war coincide on this point. They believe, as I noted, that all combatants lose their moral right not to be killed by enemy combatants in conditions of war. But what is the reason for thinking that the right they have in peacetime no longer protects them in war?

I have argued at length elsewhere against the view that just combatants forfeit their right not to be killed by posing a lethal threat to others, as well as against the view that they waive their right not to be killed.¹ If my arguments are right, we must, if we wish to preserve the traditional view, explore other possible ways of defending it. Perhaps the most promising line of argument is to be found in the commonly held view that the moral principles that govern the practice of war are different from those that govern other areas of life, including those that govern lesser forms of violent conflict, such as individual self-defense. According to this view, when war occurs, some people lose some of the rights they had in peacetime, though they may gain certain others. Yet it is hard to see why this should be so. And this view also introduces rather odd problems that would otherwise not arise. For example, it entails that it is of great moral significance to be able to distinguish accurately between wars and other forms of conflict that do not count as wars, since whether a conflict is a war determines which set of moral principles apply to it and thus which acts are morally permissible and which are impermissible.

It is, nonetheless, possible to find a basis in moral theory for such a view. It is arguable the most plausible basis can be found in the work of Hobbes. It is widely

believed that Hobbesian moral theory entails political realism, at least about war – that is, that it entails that the practice of war is not governed by morality at all, that it is outside the scope of morality altogether. This is because war occurs between states that confront one another in a state of nature, in which there is no higher authority that can coerce them to comply with rules of mutual restraint. And according to Hobbes, morality arises only when people contract to place themselves under a sovereign, and thus has no application in a state of nature. (The collectivist assumption that war is a relation between states is not essential to the derivation political realism, for individual citizens of different states at war are also in a state of nature vis-à-vis one another.)

Although political realism was influential in the US for a few decades following the Second World War, it is now widely repudiated, and rightly so. Whether the full set of Hobbesian assumptions entails political realism is an issue I will not consider. What does seem plausible is that a weakened or relaxed set of Hobbesian assumptions can support the weaker view that the moral principles governing the practice of war are quite different from those that apply in other areas of life. It seems, in particular, that a suitably weakened set of Hobbesian assumptions supports the view that the best understanding of the morality of war is closely congruent with the rules of war in the form in which we now have them, including the rules that entail the moral permissibility of the killing of just combatants by unjust combatants.

2 A Hobbesian Case for Combatant Liability and Civilian Immunity

Here is a sketch of how a Hobbesian argument for the current rules of war might be developed. Although Hobbes embraced psychological egoism and his political realist followers have assumed that states as well as individuals act only out of self-interest, we know that neither of these views is true. People generally care about whether what they do is morally permissible and are sometimes willing to sacrifice their interests in order to bring their action into conformity with their beliefs about what morality requires. This is true even of some political leaders. Most people also believe that moral constraints apply to the action of states and are concerned that their own state and its political leaders should act within those constraints.² The pressure on political leaders to act with moral restraint, even in war, tends therefore to be stronger in democratic states. But even the worst leaders in the worst states sometimes act according to certain constraints, and not always for reasons of mere prudence. The Nazis, for example, exercised restraint in the treatment of British prisoners of war, though not in the treatment of prisoners of nationalities they considered inferior, such as Russians.

People do, however, believe that their political leadership has a duty to protect and advance their interests, and they also tend to think either that the moral constraints on states are relatively weak – weaker, for example, than those that apply to the acts of individuals – or that those constraints are overridden by the national interest when the stakes become high. But even though the morally motivated demand for restraint is comparatively weak, it is reinforced by prudential concerns. People know that wars will continue to be fought and that their own state may be attacked, or start a war, or be drawn into a war by pressures from allies or for other reasons. They are aware that if their state does become embroiled in war, there will be disagreements about whether its participation is just, and that it is possible that its participation will in fact be unjust. Knowing this, people will want to be treated with restraint if war does occur and to be shielded from harsh penalties if their state's participation turns out to be unjust. They

therefore have an interest, in advance, in securing arrangements that will serve these goals.

One might doubt whether Hobbesian theory can even recognize the distinction presupposed in these remarks between wars that are just and those that are unjust. For Hobbes notoriously argues that “where there is no common Power, there is no Law: where no Law, no Injustice. Force and Fraud, are in warre the two Cardinall vertues.”³ Although this is of comparatively little significance for our purposes, since the prudential reasons that people have to try to ensure that they will be treated with restraint if war occurs do not depend on there being a distinction between just and unjust wars, it is nevertheless worth noting that there are at least two ways in which Hobbesian theory can recognize that wars can be either just or unjust. First, even if morality in domestic society is best understood, explained, and justified as a product of a Hobbesian contract, it is nevertheless a fact that many people are not Hobbesians and thus might insist that at least some of the constraints that limit what they may do to their fellow citizens also constrain what they may permissibly do to the citizens of other states, even in war. In Hobbesian terms, the social contract may commit citizens to observe certain constraints on the resort to and conduct of war even if conformity to those constraints is not *owed to* the adversaries who benefit from their observance. Second, one might argue that at least some types of war that are now widely regarded as unjust actually contravene certain Hobbesian laws of nature. An example might be a war fought for mere gain rather than to defend or preserve the state or its citizens. In *The Elements of Law*, Hobbes writes that “thus much the law of nature commandeth in war: that men satiate not the cruelty of their present passions, whereby in their own conscience they foresee no benefit to come.” This, of course, is only a very weak constraint, since it insists only that war must be sufficient for *some* good. But this passage is followed a few lines later by the broader claim that while “the want of security otherwise to maintain themselves” can justify people in harming others, nevertheless “nothing but fear can justify taking away another’s life.”⁴ To be justified, therefore, a war must at least be motivated by the desire to avoid some serious harm, not just by a desire to increase a state’s wealth or to secure other benefits or advantages.

It is rational, then, for people to want to achieve agreement with potential adversaries on rules of mutual restraint in war, and perhaps required by the laws of nature to seek such agreements. Yet international society, while not anarchic, still has no effective means of enforcing compliance with such an agreement. An effective agreement, therefore, must contain its own incentives for compliance.⁵ These are in fact not hard to supply, though like all such incentives, they are inevitably imperfect, in ways that I will explore later. What is necessary is for people to find rules of restraint that satisfy the following three conditions:

- (1) It is knowable by all to be better over time for *each* party if both comply with the rules than if neither does.
- (2) There can be no reasonable expectation that one party will continue to comply indefinitely in the absence of reciprocal compliance by the other.
- (3) It is not fatal for either party to comply for at least a limited period in the absence of immediate compliance by the other.

In these conditions, people should be able to see that their own side’s defection, or continued defection, from an agreement would eventually prompt reciprocal defection by

the adversary, which would be worse for the initial violator than continued compliance would be. Even a single instance of defection would risk the collapse of the agreement. So people's moral motives to comply would be supplemented by strong prudential incentives. (The rationality for long-term relations of this kind of strategy – that is, cooperate initially but refuse to continue to cooperate if the other does not reciprocate – has been vindicated in Robert Axelrod's well-known series of experiments with iterated Prisoners' Dilemmas.⁶)

There are various possible codes for the regulation of war that might satisfy these three conditions. Yet there are reasons why it might be rational for people to agree to accept and abide by the current rules of war rather than some alternative set of rules. And if it would be rational for potential adversaries to agree in advance to the acceptance of these rules in the reasonable expectation of reciprocal compliance, that might be sufficient, in Hobbesian terms, to justify them as *moral* rules governing the practice of war.

The central principle that requires justification is that to which I referred at the outset: the requirement of discrimination, which holds, in its conventional formulation, that whereas all combatants are legitimate targets in war, all civilians are morally immune from intentional attack. This principle seems uniquely qualified to attract agreement as a means of reducing the destructiveness of war in current conditions. This is so for a variety of reasons.

First, for any principle to have a chance of universal acceptance, it must be neutral in the sense of giving no advantage, or at least no foreseeable advantage, to any party. It cannot, therefore, discriminate between the side (if any) that fights with justification and that which does not. Any principle that sought to incorporate that distinction would in any case be doomed to ineffectiveness unless the law can provide both precisely determinate and authoritative guidance that would enable states to determine in advance and with assurance whether or not their war would be just. Otherwise the situation will continue as it is now, in which all parties to war either believe or profess to believe that they fight with justification, and in which there is no "common Power" that can authoritatively demonstrate a contrary judgment, much less enforce it. The traditional requirement of discrimination, understood as asserting universal liability to attack among combatants and universal immunity from attack among civilians, satisfies this currently necessary condition of neutrality.

Second, an acceptable principle must be not only prudentially acceptable to all parties but also morally acceptable. Potential adversaries may have developed divergent views about the morality of war within their different societies. Principles justifiable in Hobbesian terms, which are acceptable to and binding on all parties to which they apply, must abstract from those local principles. They might be formed from points of convergence, or they might constitute mutually acceptable compromises. The requirement of discrimination has been promulgated for centuries by theorists in the just war tradition, is also enshrined in international law, and seems to constitute a point of overlap among the teachings about the morality of war of the major religions. It is therefore now about as uncontroversial among the peoples of the world as any principle for the limitation of violence in war could be.

The requirement of discrimination contains both a permission and a prohibition. It permits the combatants of all parties to a war to kill combatants on the opposing side, or

sides. It must do this to have any chance of acceptance. For no people with a concern for their own interests could accept a principle for the conduct of war that might forbid them to fight at all, particularly when they believe, though wrongly, that fighting is justified. Yet the permission to kill opposing combatants is a highly significant permission: it exempts those who fight without justification from any liability for what might otherwise count as murder, provided that they adhere to the rules of *jus in bello*. The permission is, moreover, logically related to another benefit that the current rules of war confer on combatants on all sides – namely prisoner of war status in the event of capture. When all parties accept that it is permissible for all combatants, including their own enemies, to kill combatants on the opposing side, they have no grounds for punishing captured enemy combatants who have fought by the rules or for doing any more to them than detaining them until the war is concluded.

The permission granted by the requirement of discrimination may, indeed, be seen as a reward offered to combatants on both sides for respecting the prohibition. It is as if combatants are offered immunity from any of the ordinary penalties for killing other people, as well as the benefits of prisoner of war status, in exchange for refraining from attacking and killing civilians.

The advantages that the prohibition of the intentional killing of civilians offers to all belligerents are numerous and obvious. Perhaps most importantly, general respect for civilian immunity has the effect of insulating ordinary life from the direct effects of war, thereby enabling civilized life to continue even in the midst of war. Some just war theorists have, indeed, understood the importance of civilian immunity primarily in collectivist terms. Michael Walzer, for example, contends that “the deepest meaning of noncombatant immunity” is that “it doesn’t only protect individual noncombatants; it also protects the group to which they belong.”⁷

A third advantage of the principle of civilian immunity is that the protected category is definable with comparative precision. Although there are gray-area cases, the distinction between combatants and noncombatants (and the closely related but not identical category of civilians) is perhaps as precise and unambiguous a distinction as one can hope for in this area.⁸ It is certainly easier to apply than, for example, the distinction between those who are and those who are not responsible for threats of wrongful harm, which some philosophers have argued corresponds to the distinction between morally legitimate and illegitimate targets in war.

Two other virtues of the principle of civilian immunity are worth mentioning. One is that, unlike any constraint on attacking enemy combatants, it does not restrict a state’s capacity for self-defense. This is because civilians generally, and almost by definition, do not pose a threat and therefore in general cannot be a target of action that is in the literal sense defensive. The killing of civilians can of course serve certain purposes in war – for example, it can bring pressure on a government to surrender, as the American nuclear bombings of Japanese cities were intended to do at the close of the Second World War. But because this sort of terrorist action does not operate directly, the way defensive action does, but must operate through the wills of people other than those attacked, it is of notoriously uncertain effectiveness, as Osama bin Laden and his confederates discovered when they sought to eliminate the American military presence in the Middle East by attacking the World Trade Towers in Manhattan. Hence it is reasonable for a state to expect that compliance with the principle will not threaten its security.

Finally, because the principle of civilian immunity has long been hallowed by tradition, a considerable motive for compliance with it already exists: namely, conformity with tradition.

That there are these many reasons for rational convergence on the requirement of discrimination, with its constituent principle of civilian immunity, strongly supports the Hobbesian argument for the claim that this centerpiece of the current rules of war is indeed a requirement of morality, not just an artifact of positive law. The worry about this Hobbesian argument, however, is that it presupposes the Hobbesian understanding of moral justification, which grounds morality in what largely self-interested people as they actually are (that is, without any restrictions, such as Rawls's "veil of ignorance," on what they can know) could rationally agree to. On this view, we have to understand the moral reasons we have not to kill people as the products of rational agreement motivated largely by self-interest. That the moral constraints on killing have this foundation is essential to the Hobbesian explanation of why the reason we have not to kill innocent people within our own society does not extend to the killing of just combatants in war, however innocent they may be; for in those altered conditions, a different kind of agreement becomes rational.

This worry is not merely theoretical. We can make the implications of the Hobbesian argument vivid by reflecting that the conditions I have described in which it seems rational for states to agree to adhere to the requirement of discrimination as traditionally interpreted do not necessarily characterize the situation of conspicuously weak states, much less that of even weaker nonstate groups with political grievances. In current conditions, it would *not* be self-interestedly rational for a weak, nonstate group to agree to be bound by the principle of civilian immunity. This is so for two familiar reasons. First, such groups have no chance of achieving their aims by conventional military means against the more numerous and heavily-armed forces of an adversary state. The killing of civilians may be their only effective possibility for the use of violence in achieving their aims. Second, they generally need not fear retaliation in kind, in part because they can generally rely on moral pressures, and pressures of public relations, to deter states from engaging in terrorism, but even more because they tend not to be representative of any precisely identifiable group that could be discriminately targeted in a terrorist reprisal. When these conditions obtain, it would not be rational in Hobbesian terms for a small group to agree to abide by the principle of civilian immunity; therefore the group has no moral reason, again in Hobbesian terms, not to pursue their ends by terrorist means.

I find this a fatal defect in this Hobbesian argument for the claim that the rules of war as we have them today are constitutive of the morality of war in current conditions. But one might, I suppose, continue to hope that it can be shown that the argument does not in fact imply that the requirement of discrimination is not binding on small, nonstate political groups. There is, however, a further problem, which is that this Hobbesian argument for the current rules of war seems incompatible with other central elements of Hobbes's moral and political philosophy – elements that are identified and elaborated in Sharon Lloyd's contribution to *Hobbes Today*. Lloyd appeals to these other elements in an effort to provide a different Hobbesian vindication of the requirement of discrimination as traditionally understood. I will explain at the end of the paper why, if Lloyd's understanding of Hobbes is right, it undercuts the argument I have given.

3 Lloyd's Arguments

Lloyd argues that Hobbesian moral and political theory has the resources to justify both the prohibition and the permission that together constitute the requirement of discrimination as it is traditionally and almost universally understood.⁹ Her Hobbesian case for the prohibition of the intentional killing of civilians – the principle of civilian immunity – takes the form of a response to those who have argued that some civilians can be liable to suffer certain harms in war if they share responsibility for the wrong that the war is fought to prevent or correct. Although she is sympathetic to the idea that responsibility for a wrong is a basis of liability to be harmed by action that would prevent that wrong, Lloyd argues that civilians are not responsible for threats of wrongful harm that their country may pose in war. One might think that Hobbes would find them responsible because their sovereign acts on their authorization. But in fact the civilian citizens cannot authorize their government to fight an unjustified war of aggression because no citizen has the right to contribute to such a war, since it is prohibited by the law of nature, and one cannot authorize another to do on one's behalf what one has no right to do oneself. If the sovereign initiates such a war, it acts without the citizens' authorization. They are, nonetheless, obligated to obey its command to support the war. And because they are morally obligated to support the war, they cannot be morally responsible for doing so. Responsibility for the war is the sovereign's alone. Hence only the sovereign, not the citizens, can be liable to suffer the effects of defensive or corrective action.

The crucial claim here is of course that the citizens are morally required to obey the command of a legitimate authority to support an unjust war. Lloyd identifies three Hobbesian bases for this requirement. First, the law of nature "that forbiddeth breach of covenant" overrides other laws of nature, including laws that prohibit supporting or contributing to an unjust war.¹⁰ Second, given the pervasiveness of disagreement about the morality of any particular war as well as the vagaries and general unreliability of private judgment about such matters, citizens will be more likely to act in conformity with the laws of nature on more occasions over time if they defer to the authority of the sovereign than if they act on the basis of private judgment. As an analogy, Lloyd cites the fact that children are more likely to remain unharmed if they obey a responsible, though fallible, babysitter than if they are guided by their own judgment. Third, Lloyd claims that "the only alternative to joint submission to a public authority consistent with the reciprocity requirement of the Law of Nature is a state of universal private judgment." If this is right, it seems that universal submission is necessary for coordinated collective action, which would be largely impossible if everyone acted on the basis of their private judgment.

The most important objection to this defense of civilian immunity is related to Lloyd's case for the permissibility of fighting even on the unjust side in war and I will therefore reserve this objection until near the end of the paper. There are, however, other objections. One is that there *are* ways in which civilians could bear responsibility for an unjust war even on Hobbesian assumptions. Suppose, for example, that the question of whether to go to war has to be addressed and the sovereign authority decides to resolve the question by plebiscite. It announces that the state will resort to war if, but only if, a majority of the citizens votes in favor of that option. Suppose that a majority then votes to go to war and an unjustified war of aggression ensues. Even if the citizens cannot

validly authorize the sovereign to pursue an unjust war, and even if they are obligated to support the war once it is declared by the sovereign, it is hard to deny that those who voted for it bear some moral responsibility for it, given that it would not have been fought had they not freely voted for it. This is not, however, a significant objection in practice, since unjust wars are in fact never initiated in this way.

A second objection is that it seems implausible to suppose that the prohibition of violating covenants (that is, of breaking promises or breaching contracts) always overrides other prohibitions, such as the prohibition of intentionally killing innocent people. And the claim that people will do better over time in conforming their conduct to the laws of nature if they obey the sovereign rather than acting on their private judgment seems equally implausible as a universal claim. It may be true of most people in relation to a particularly trustworthy sovereign, or political authority, but false of most people in relation to a sovereign, such as the Nazi government in Germany, that is either blind or indifferent to the laws of nature. (In these latter cases, the relation between the sovereign and the citizens is less like the relation between a conscientious babysitter and the children in her charge and more like that between Fagin, in *Oliver Twist*, and the children he controls.) And even in the case of a morally scrupulous sovereign, there may be many citizens whose private judgments are generally more reliable in discerning the demands of the laws of nature than the sovereign is. Of course, if it is true, as Lloyd claims, that the only alternatives consistent with the requirement of reciprocity are universal submission and universal reliance on private judgment, then there cannot be a rule that binds the morally unenlightened to universal submission while allowing the morally enlightened an occasional license to act on private judgment. But it seems consistent with the requirement of reciprocity to require general though not exceptionless submission by everyone while also allowing everyone a restricted right of reliance on private judgment in certain cases. This might be more likely to promote maximal compliance with the laws of nature, even under a conscientious sovereign, than universal submission to the sovereign, and could also be compatible with the possibility of coordinated collective action. That there can be a limited right of epistemic self-reliance even in opposition to the commands of the sovereign seems implied by Lloyd's understanding of why Hobbes asserts that a judicially guilty person may defend his life when threatened with execution. According to Lloyd, the explanation is not that the law of nature grants an exceptionless right of self-preservation but because the convicted criminal is not obliged to accept the judgment of the sovereign, expressed through the judicial system, that he is guilty. (Her further claim that the criminal is blameless for acting in a way "that we may not reliably be able to resist" asserts only an excuse rather than a right or permission.)

A further reason for doubting that citizens are more likely to obey the laws of nature, and therefore to flourish to a greater degree, if they consistently defer to the judgment of their sovereign is that, at least with respect to war, their sovereign's judgment will be contradicted by that of the enemy sovereign. Unless they have good reason to believe that their own sovereign is epistemically more reliable than the enemy sovereign, it seems arbitrary for them to defer to the judgment of their own sovereign. But if their epistemic deference to their own sovereign is based on their belief that their sovereign is more reliable epistemically than the enemy sovereign (and than they are), it seems that they are ultimately relying on private judgment after all – not their judgment

about individual cases but their judgment of comparative epistemic reliability. If, however, they have no good reason to believe that their own sovereign is a more reliable source of judgment than the enemy sovereign, it seems that they might do just as well in conforming to the laws of nature by universally and consistently submitting to the judgment of the enemy sovereign as by submitting to the judgment of their own. They might thereby even do *better* if, for example, they are citizens of a totalitarian state while the enemy state is a democracy with a free press. Unless citizens have good reason to believe that some sovereign is worthy of epistemic deference, the most they can confidently expect to achieve through universal submission to *any* sovereign is not enhanced *reliability* of judgment, and therefore a higher probability of conformity with the laws of nature, but mere epistemic *coordination* and therefore *uniformity* of judgment – no doubt an advantage of sorts but not the one advertised.

Consider now Lloyd's Hobbesian case for the permissive component of the requirement of discrimination: the right of combatants, including unjust combatants, to attack and kill enemy combatants. One suggestion is that the permission that unjust combatants allegedly have to kill just combatants is entailed by the right of epistemic self-reliance in cases of factual or normative disagreement. If there is such a right, it may indeed make it permissible in what Parfit called the "belief-relative" sense, or even the "evidence-relative" sense, for *some* unjust combatants to kill just combatants.¹¹ That is, some unjust combatants' action will be permissible *relative to* their beliefs or to the evidence accessible to them. If, however, their beliefs are unreasonable, they may be culpable even though there is a sense in which their action is permissible. And even if their false beliefs are reasonable, so that their action is not culpable, that action can nevertheless, on some views, be a basis of liability if it is objectively wrong and they are aware, or ought to be aware, of the risk that it might be so. But if many unjust combatants are liable to attack while just combatants are not, that is a significant moral asymmetry.

While Lloyd is thus correct to say that some unjust combatants are permitted to kill just combatants, this is true only in the belief-relative or evidence-relative sense of permissible. And this may not be sufficient to shield them from liability. There are moreover, some unjust combatants who lack even the belief-relative permission to kill just combatants – namely, those who know or strongly suspect that their war is unjust in the sense that it contravenes the laws of nature. The argument from the right of epistemic self-reliance does nothing to justify their action or to relieve them of liability.

One may also question whether unjust combatants actually do have a right of epistemic self-reliance. While the Hobbesian view clearly implies that persons involved in individual conflicts have that right, it seems that soldiers are under a requirement of epistemic *deference* to the sovereign. But, as we have just seen, the principal reason that Lloyd cites for epistemic deference to the sovereign may in some cases support deference to the judgments of the enemy sovereign rather than to those of one's own.

It might be better, therefore, to rest the Hobbesian case for the permission that unjust combatants allegedly have to kill just combatants on the same ground to which Lloyd appeals in rejecting civilian liability: namely, the duty of obedience to the sovereign. Recall that according to Lloyd's reading of Hobbes, this duty of obedience, based on the prohibition of the breach of covenants, overrides other laws of nature. Thus it is compatible with its being otherwise wrong for unjust combatants to kill just

combatants for it to become obligatory for them to do so under command of their sovereign. As with civilians who support an unjust war, the fact that unjust combatants act under an obligation when they kill just combatants frees them of responsibility, which lies only with the sovereign. Yet, as Lloyd points out, if they are not responsible and liability is grounded in responsibility, they cannot be liable to attack. And the same is true of just combatants. This leaves us with a view of war according to which *no* combatants are morally liable to attack, so that the justification for killing in war might have to derive simply from the duty of obedience to the sovereign.

There is, however, a problem with the idea that combatants, and in particular unjust combatants, are obligated to kill in war by the command of their sovereign. Assume, as Lloyd suggests, that some wars are contrary to the laws of nature and that participation in them is thus forbidden by those same laws. Yet Lloyd claims that participation in such wars is nevertheless obligatory if commanded by the sovereign. Yet it is questionable whether there is actually an obligation of obedience in these cases. The problem is Hobbes's claim, quoted by Lloyd, that "unless he that is the author hath the right of acting himself, the actor hath no authority to act." I know of no ground for supposing that the sovereign has a right to violate the laws of nature. He cannot have it in his own person independently of the powers he acquires as sovereign, and he cannot derive it by transfer from his subjects, since they too lack it. But if the sovereign has no right to violate the laws of nature, the claim just quoted implies that he has no power to authorize citizens to violate them. The command of the sovereign, therefore, cannot authorize combatants to fight in a war that is in violation of the laws of nature. The laws of nature seem not to be self-effacing in the way that Lloyd suggests they are.

It may well be, however, that Hobbes or contemporary Hobbesians have the resources to avoid this apparent inconsistency. So assume for the sake of argument that the sovereign does have the power to make it obligatory for its citizens to fight in or otherwise support a war that is in violation of the laws of nature – an assumption that is also necessary for the success of Lloyd's argument against civilian liability. On this assumption, combatants who fight at the command of their sovereign cannot be liable for doing so, for all responsibility for their action lies with the sovereign. Lloyd then locates a source of the permission to attack them that is independent of liability and also, perhaps, rebuts the charge that they act in violation of the laws of nature, even when they fight in a war that we would identify as unjust. This is what Lloyd and others refer to as the Hobbesian "soldier contract," according to which soldiers who enlist in the military consent to accept "an obligation to be the target of enemy hostilities." This confers a permission on enemy combatants to attack them, so that the permissibility of attacking combatants in war derives not from their liability but from their consent.

There are four problems with this view. One is that it provides no basis for the permissibility of attacking conscripts, who, as Lloyd concedes, have not voluntarily accepted an obligation to be objects of attack. Thus, on Lloyd's view, an army composed entirely of conscripts that fights for a just cause would be permitted to attack enemy volunteers, but those volunteers would have no moral basis for fighting either in the liability or the consent of their conscript adversaries. The moral equality of combatants cannot, then, be defended as a *universal* moral doctrine on the basis of an appeal to the soldier contract.

The second problem with the soldier contract is that it is a contract *within* the soldier's own political community that commits him to his fellow citizens to assume the risks of combat. But from the fact that volunteer soldiers have taken on "an obligation to be the target of enemy hostilities" it does not follow that "it becomes morally permissible to target them." A contract with one's fellow citizens to assume the risks of combat does not confer any sort of permission on one's adversaries.

But suppose that it did involve granting such a permission. Suppose that the soldier contract somehow involves consenting to be a *legitimate* target of attack. The third problem is that, even on that assumption, it is doubtful that such a free expression of consent could actually succeed, *on its own*, in making an attack legitimate. I might, for example, freely grant you permission to kill me now, yet few would accept that that alone could make it permissible for you to kill me.

Finally, even if unjust combatants were permitted to kill just combatants because of the latter's consent, it still would not follow that it is permissible for them to fight in an unjust war. For participation in an unjust war might be wrong because of what it involves doing to enemy people generally, and not just to enemy combatants. It might be wrong because it would, in Lloyd's words, "violate the fundamental reciprocity requirement of the Law of Nature, that they not behave in a way that they would condemn in others."

It seems, therefore, that Hobbesian theory fails to ground the independent moral permissibility of the killing of just combatants by unjust combatants. The only justification seems to derive from the obligation of obedience to the sovereign, which overrides any law of nature that prohibits participation in an unjust war. So the case for the claim that unjust combatants are justified in killing just combatants seems to depend, in the end, on the claim that the sovereign can effectively authorize citizens to act contrary to the laws of nature and that this authorization can create an overriding obligation that exempts citizens, whether combatants or civilians, from responsibility and therefore from liability for their contribution to a war that violates the laws of nature.

But if the sovereign can obligate citizens to act in ways that violate the laws of nature, thereby relieving them of liability for doing so, this ultimately vitiates the practical significance of Lloyd's argument against civilian liability and, more importantly, subverts the rationality of a contract among states, of the sort I described earlier, to adhere to a principle of civilian immunity as a constraint on the destructiveness of war. To see why this is so, consider what the overriding duty of obedience to the sovereign implies about the terrorist killing of civilians. If the sovereign political authority commands its soldiers to conduct terrorist attacks against innocent civilians, the soldiers are obligated to obey. According to Lloyd, the reciprocity requirement forbids us to hold the soldiers liable for merely fulfilling their obligation to submit to their sovereign. Soldiers – even unjust combatants – who engage in terrorist action against innocent civilians under the command of their sovereign are not, therefore, liable to defensive action for doing so.

If the soldier contract actually made all volunteer soldiers legitimate targets, there would be no need for a doctrine of combatant liability in order to justify defensive action against volunteer soldiers engaged in terrorism, for killing them would be justified by their having consented to be killed. (Note, however, that the justification for the defensive killing of volunteer unjust combatants who engage in terrorism would be just

the same as the justification for killing volunteer just combatants who scrupulously confine their attacks to military targets, and this seems implausible.) But the soldier contract does not legitimize the killing of volunteers and, even if it did, there would be no permission grounded in Hobbesian assumptions for defensive action against *conscripts* engaged in terrorist action under command of their sovereign.

Lloyd's argument for civilian immunity in war is thus in an important sense self-undermining. The obligation of obedience to the sovereign may exempt civilians from liability to attack, but same obligation of obedience justifies combatants in killing civilians if they are commanded to do so by the sovereign. The civilians' moral immunity is no barrier to the combatants' action, since the obligation of obedience overrides other laws of nature, including any that might require respect for civilian immunity.

The problems do not end here. The categorical nature of the obligation of obedience not only undermines the practical significance of civilian immunity but also threatens the Hobbesian argument I gave earlier for the rationality among potential belligerents (except for members of weak and unrepresentative political groups) of agreeing to abide by the principle of civilian immunity even in the absence of a global sovereign who could enforce such an agreement. That argument was premised on the assumption that people on all sides can reliably expect their potential adversaries to be motivated by the recognition that compliance with the principle will serve their own interests better than noncompliance. But if the obligation of obedience is absolute and it is reasonable to expect combatants to fulfill that obligation (as in fact they almost invariably do), then whether enemy combatants will comply with the principle of civilian immunity will be largely unaffected by their *own* perception of what it would be rational for their side to do. Their action will instead be guided by the commands of their sovereign political authority, particularly when obedience exempts them from liability. But this makes the expectation of compliance by one's adversaries much more precarious. For whether they will comply depends almost exclusively on the beliefs and motives of their political leaders. And their leaders may not be rational, or may have interests that conflict with those of the citizens. They might, for example, mistakenly believe that their adversaries will persist in adhering to the principle of civilian immunity unilaterally. Or a more likely possibility is that even if the leaders are rational, they may be willing to gamble with the lives of their civilian citizens in an effort to reap certain benefits for themselves. History is littered with political leaders who have cared very little, other than for instrumental reasons, about the lives of the great mass of their citizens. Witness Saddam Hussein's persistent refusal to comply with the UN's demand for verification that Iraq did not possess weapons of mass destruction, even when it did not in fact possess them and the refusal to permit verification resulted in the continued imposition of economic sanctions that were disastrous for ordinary Iraqi citizens. Another possibility is that political leaders may believe that the breakdown of reciprocity will not be worse for their own citizens, either because the adversary lacks the capacity to inflict significant harm on their civilians (as was the case with US leaders when they ordered the obliteration of the cities of Tokyo, Hiroshima, and Nagasaki), or, perhaps, because they believe that martyrdom will actually be good for their citizens, who will be rewarded in an afterlife.

Given that some political leaders are more likely to take irrational risks with the lives of their citizens than those citizens themselves would be, and that some political leaders are willing to sacrifice their citizens' lives for their own personal gain, it becomes less reasonable to expect all parties to comply with the principle of civilian immunity when the decision about compliance will be made by political leaders rather than by citizens themselves. Political leaders may simply be less sensitive than their citizens to the fact that it is better for the citizens of each state if their own state complies with the principle of civilian immunity. And the less reason one has to expect compliance by one's adversary, the less reason one has, on Hobbesian assumptions, to comply oneself.

The dilemma here is that the contract within each society that establishes an overriding obligation of obedience to the commands of the political authority may undermine the rationality of agreeing to a contract among societies to accept the principle of civilian immunity as a constraint on the destructiveness of war. The domestic contract subverts the international contract. It may simply not be rational to bind oneself to adhere to the principle of civilian immunity when the only ground that one can have to expect that one's adversaries will reciprocate is the hope that their political leaders will be sufficiently rational and sufficiently concerned to protect their own civilian populations to do so.

It seems, therefore, that despite Lloyd's valiant, ingenious, and resourceful efforts, even Hobbes may be unable to rescue the central elements of the orthodox account of the just war.¹²

Jeff McMahan

¹ Jeff McMahan, *Killing in War* (Oxford: Clarendon Press, 2009), sections 1.2 and 2.2.

² This apparent fact challenges Chris Naticchia's claim that citizens may be presumed to have authorized the sovereign to pursue only the national interest, and that its pursuit of goals that conflict with the national interest must therefore constitute violations of the authority granted to it by the citizens. See Chris Naticchia, "Hobbesian Realism: A Reappraisal," *Hobbes Today...* [Need page numbers from the proofs.]

³ Thomas Hobbes, *Leviathan*, chapter 13. [Need full citation.]

⁴ Thomas Hobbes, *The Elements of Law Natural and Politic: Human Nature and De Corpore Politico*, J.C.A. Gaskin, ed., Oxford University Press, 1994, p 104.)

⁵ Since drafting this paper, I have become aware of a much more detailed and sophisticated argument along similar lines in Claire Finkelstein, "Rational Contractarianism and International Law," which is a draft of chapter 11 in her forthcoming book, *Hobbesian Theory of Law*.

⁶ Robert Axelrod, *The Evolution of Cooperation* (New York: Basic Books, 1984).

⁷ Michael Walzer, "Terrorism and Just War," *Philosophia* 34 (2006): 3-12, p. 4.

⁸ For a recent challenge to traditional assumptions about the distinction between combatants and noncombatants, see Cécile Fabre, "Guns, Food, and Liability to Attack in War," *Ethics* 120 (2009): 36-63.

⁹ S.A. Lloyd, "Democratic Sovereignty and Moral Liability," *Hobbes Today...* Add citations here and elsewhere at the proof stage.

¹⁰ See the passage that Lloyd quotes, with some substitution of terms, from chapter 16 of *Leviathan*.

¹¹ See Derek Parfit, *On What Matters*, forthcoming.

¹² I am very grateful to Jerry Gaus for written comments on an earlier draft of this paper, and to Sharon Lloyd and Claire Finkelstein for much illuminating discussion. All three have just cause to lament my obtuseness in Hobbes scholarship.