Leiden Journal of International Law (2018), 31, pp. 171–194
 doi:10.1017/S0922156517000498

INTERNATIONAL CRIMINAL COURTS AND TRIBUNALS

Revisionist Just War Theory and the Concept of War Crimes

ALEJANDRO CHEHTMAN*

Abstract

Under contemporary international law, war crimes are conceived as particularly serious violations of the laws of armed conflict. Mere participation of rank-and-file soldiers in an unjust or unlawful war is generally not considered to warrant legal punishment. This position is based on the principle of equality between belligerents. During the last 20 years, this principle has been challenged by the so-called revisionist position in just war theory, as well as by certain scholars in international law. According to them, unjust or unlawful participants in armed conflict perpetrate serious wrongs. This article argues that their conduct is not only morally wrongful, but also that it should be criminalized under certain circumstances. On the basis of empirical research on cognitive biases, and on one of the leading accounts of legitimate authority in political philosophy, it argues that participation in war warrants criminalization only when the war is knowingly or manifestly unlawful. Furthermore, it claims that this position is not only sound at the level of deep moral principles, but that in fact it provides a persuasive reinterpretation of existing international law.

Keywords

international crimes; just war theory; terrorism; unprivileged combatants; war crimes

1. INTRODUCTION

Under contemporary international criminal law, war crimes are standardly conceived as serious violations of the laws of armed conflict which carry individual criminal liability. They include wrongful acts such as torture, taking of hostages, rape, sexual slavery and other forms of sexual violence, provided that they have a sufficient link with an armed conflict. They also involve conduct in violation

* Associate Professor of Law, Universidad Torcuato Di Tella [achehtman@utdt.edu]. I am grateful to the participants in the ASIL Interest Group on International Criminal Law workshop at NYU, Seminario Carlos Nino at SADAF (Argentina), the War in the 21st Century Conference at Di Tella Law School, and the Answering for International Crimes Workshop at Pluricourts (Oslo), especially Margaret deGuzman, Adil Haque, Ryan Liss, Anni Pues, and Alain Zysset. I am also indebted to Maximo Langer for useful discussion and to Cécile Fabre, Charles Garraway, Massimo Renzo, Gerry Simpson, Kathryn Sikkink and three anonymous reviewers for detailed written comments and suggestions. All websites were last visited on 20 September 2017. This work has been funded by a Marie Curie Fellowship of the European Commission and a PIP Grant from the Argentine National Research Council (CONICET). The usual disclaimer applies.

of the cardinal principles of the laws of armed conflict, such as the principles of discrimination, necessity, and proportionality. Accordingly, the Rome Statute provides criminal liability, *inter alia*, for those who intentionally direct attacks against the civilian population as such or against civilians not taking direct part in hostilities, and those who intentionally launch an attack in the knowledge that such attack will cause incidental loss of life or injury to civilians which would be clearly excessive in relation to the concrete and direct overall military advantage anticipated.[1] A structural feature of the laws on war crimes is that they standardly apply symmetrically between *ad bellum* lawful and unlawful belligerents. This approach was endorsed and advocated by the so-called orthodox position in just war theory for the greater part of the twentieth century.[2]

Nevertheless, contemporary just war theory has undergone a significant revolution over the last two decades. Following the classical works of Vitoria, Suárez, and Grotius, the so-called revisionist or neo-classical approach has challenged the accepted view that just and unjust belligerents stand in morally symmetrical positions.[3] Unlike those fighting a just war, participants in an unjust war are typically liable to being killed by the just party. This analysis has deep implications for the notion of criminal conduct in war.[4] While both orthodox and revisionist just war theorists consider certain types of behaviors – such as intentionally killing civilians or torturing prisoners of war – as war crimes irrespective of who perpetrates them, they seem bound to diverge on a particularly important issue: unlike orthodox theorists, the revisionist approach argues that mere participation in an unjust war is morally wrongful, often described as on par with participating in murder.[5]

This position may seem to entail a significant departure from the laws on war crimes as they currently stand. Yet, it is increasingly relevant and indeed appealing in a world of asymmetrical conflicts, with their trend to more ubiquitous violence, and of extensive movement of people across borders.[6] Consider, for instance, the case of a former Islamic State of Iraq and the Levant (hereinafter, ISIS) fighter that arrives at the airport in Brussels, or a recruiter working in the outskirts of Madrid. Are their acts not morally reprehensible to the extent that they should warrant criminal sanctions? Should national authorities not be entitled to

[1] Statute of the International Criminal Court, 2187 UNTS 90, Arts. 8(2)(b)(i), (iv), (xx), and 8(2)(e)(i) (hereinafter, Rome Statute) although many other war crimes can ultimately be traced to the violation of one or more of these principles.

[2] See, e.g., M. Walzer, *Just and Unjust Wars* (1992).

[3] J. McMahan, *Killing in War* (2009); D. Rodin, *War and Self-Defense* (2002); and C. Fabre, *Cosmopolitan War* (2012).

[4] A point of clarification is in order here. I will not provide an account of when or why a particular moral wrong should be criminalized. This is a thorny question in domestic criminal law theory which does not seem to be critical to our enquiry here, given that most of the grave wrongs we shall consider certainly satisfy the requirements for criminalization under almost any existing approach – the question, by contrast, is about the conditions under which they should be criminalized.

[5] See Section 2, *infra*.

[6] See, generally, M.L. Gross, *Moral Dilemmas of Modern War: Torture, Assassination, and Blackmail in an Age of Asymmetric Conflict* (2010).

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

prosecute and punish them for participating in this type of heinous enterprise? And, if so, should these individuals not be held accountable under domestic penal regulations or under international law? Currently, this type of conduct is being prosecuted under domestic law as ordinary criminal offences, often as participation in terrorist activity.[7] Yet, often no conduct was performed on the territory of the prosecuting state, by or against its nationals, or even necessarily against its sovereign interest, which constitute the standard bases for domestic criminal jurisdiction.[8]

This article argues, by contrast, that it would be normatively adequate for domestic authorities (and arguably also for international courts) to call these individuals to account under international law as war criminals. And that this would be the case even if the individual in question has not been personally involved in the intentional execution of innocent civilians, the perpetration of sexual crimes, torture, or other 'traditional' war crimes. I argue that it suffices that they contribute to such a monstrous enterprise to consider them morally responsible, and indeed criminally liable, for serious wrongdoing.

Accordingly, I advocate two main propositions. First, I articulate a particular version of the revisionist account which suggests that the underlying acts that constitute participation in an unjust war warrant liability to legal punishment, albeit only when the war is knowingly or manifestly unjust. This limited thesis is not only compatible with the central insight advocated by revisionist theorists – in fact, it follows closely the positions of Suárez, Vitoria and Grotius – but is also better attuned to our basic understanding of what amounts to criminal behavior in general.[9] Second, I argue that this particular moral position, far from requiring a radical modification to the existing law, is supported by the current legal framework to a larger extent than alternative orthodox and revisionist accounts. In fact, it advocates an innovative interpretation of this framework, and proposes minor adjustments to make it still more attuned to our basic moral convictions.

The article proceeds as follows. Section 2 spells out the specifics of the revisionist position in just war theory and its main implications for the identification of morally wrongful conduct in war. In Section 3, I examine when these wrongful conducts warrant criminalization at the bar of justice. I criticize Cecile Fabre's radical revisionist position on war crimes for being over-inclusive and I present my own revisionist argument based on both empirical and normative considerations. Section 4, in turn, addresses six objections to this position raised by Jeff McMahan. In Section 5, I explain how the normative argument advocated in this article is to a significant extent compatible with the way in which criminal responsibility for wrongful conduct in war is regulated as a matter of law. I also advocate certain amendments to the existing laws in order for them to better capture the underlying moral principles. Section 6 briefly concludes.

[7] See Section 5, *infra*.
[8] See A. Chehtman, 'Jurisdiction', in M.D. Dubber and T. Hörnle (eds.), *The Oxford Handbook of Criminal Law* (2014), 399.
[9] See Section 3, *infra*.

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

2. The revisionist position and wrongful conduct in war

Revisionist just war theorists argue that there is continuity between everyday morality and the morality of war.[10] War situations are interpersonal situations writ large – the main difference is the number of persons involved and the degree to which their action is co-ordinated. Accordingly, much as in interpersonal situations, in which the positions of a culpable attacker and her victim are morally asymmetrical, so must be the position of just and unjust belligerents.[11] For simplicity, I shall stipulate here that any just war must meet the following requirements: (i) it has a just cause, where a just cause consists of the violation, backed by the threat of lethal force, of some party's fundamental human rights; (ii) it is a proportionate response to the injustice that the belligerent is suffering or is about to suffer; (iii) it is not fought and won through the deliberate and indiscriminate targeting of innocent non-combatants; (iv) it stands a reasonable chance of succeeding by military means that do not breach the *in bello* requirements of proportionality and discrimination; and (v) there is no less harmful way to pursue the just cause (*ultima ratio*).[12]

This point of departure leads revisionist just war theorists to challenge the principle of separation between *ad bellum* and *in bello* considerations. They typically argue that *ad bellum* considerations have deep implications for the appropriate *in bello* rules.[13] Yet, an often-underappreciated aspect of this influence is that it cuts both ways. As indicated, one of the necessary conditions for a just war is often believed to be that it is not fought and won through the deliberate and indiscriminate targeting of innocent non-combatants (as per (iii) above), i.e., that it complies with fundamental *in bello* principles.[14] Accordingly, I will refer here to an unjust war as one which fails to satisfy either the *ad bellum* or *in bello* requirements, or both.

The revisionist position, thus described, has far-reaching implications for what constitutes wrongful conduct in war. As Cheyney Ryan recalls, Augustine's writings already suggest that '[s]oldiers fighting an unjust war were akin to criminals, who had no more right to commit violent acts than an ordinary criminal [while] soldiers on the just side were like magistrates bringing the criminal to justice'.[15] This means that, whereas those fighting on the just side commit no wrong by killing unjust (liable) soldiers (in fact, they would often be acting with positive justification), those

[10] See, e.g., McMahan, *supra* note 3.

[11] This is subject to a few qualifications. Some of those fighting an unjust war whose responsibility is greatly mitigated by ignorance or duress, and who will contribute very little or nothing at all to the unjust aims, would not be liable to being killed. By contrast, some of those fighting with just cause, but who do so by unjust means, or exceed their right to defend themselves, may also become liable. I will not address this further issue here.

[12] I follow here Fabre, *supra* note 3. See, similarly, T. Hurka, 'Proportionality in the Morality of War', (2004) 33(1) *Philosophy and Public Affairs* 34, Section 1; McMahan, *supra* note 3.

[13] McMahan, *supra* note 3, at 203 and accompanying footnotes.

[14] See, e.g., C. Fabre, *Cosmopolitan Peace* (2016), 7.3.

[15] C. Ryan, 'Democratic Duty and the Moral Dilemmas of Soldiers', 2011 122(1) *Ethics* 10, at 15.

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

fighting an unjust war would be violating the rights of those fighting on the just side by killing them.[16]

This insight has substantial implications for the cardinal *in bello* principles, namely, distinction, necessity, and proportionality. Distinction is often taken to entail a flat prohibition of deliberately attacking civilians and civilian objects, or a requirement that attacks are only directed to legitimate targets.[17] Accordingly, the principle of distinction grounds a stringent moral requirement not to intentionally harm those who are not liable to being targeted, but it considers it morally permissible to attack those who are liable. The key question here is who is liable to being attacked in war. For revisionist just war theorists, liability does not depend on whether fighter A formally belongs to the army of belligerent X, but rather on fighter A being responsible for an objectively wrongful threat of serious harm.[18] As a result, the revisionist understanding of the principle of distinction would allow harming those fighting an unjust war but would prohibit harming those on the just side.[19]

Second, the principle of necessity is similarly affected. At the very least, necessity rules out any measure which is not indispensable for securing the ends of the war.[20] However, if *per hypothesis* the war aims are morally unjust, then it follows that any act that is necessary to achieve those aims would also have to be morally impermissible. That is, harms that are necessary only for the purposes of obtaining a morally wrongful goal cannot be morally justified. Accordingly, revisionist just war theory suggests that a more accurate understanding of the principle of military necessity would assert:

> that an act of war that would harm innocent people as a side effect is impermissible if there is an alternative act that would cause less harm to innocent people and make an equal or greater contribution to the achievement of a just or justifying aim.[21]

As a result of this, whereas acts of the just belligerent *can* satisfy the requirement of necessity, acts of the unjust one typically *cannot*.

---

[16] See, e.g., S. Neff, *War and the Law of Nations* (2005), 63, citing F. Suárez, *A Work on the Three Theological Virtues: Faith, Hope, and Charity* (1958), 813.

[17] See *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion of 8 July 1996, [1996] ICJ Rep. 226, para. 78; Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977, 11 UNTS 3, Arts. 48, 51(2), and 52(2) (hereinafter, Additional Protocol I or AP I); Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II), 8 June 1977, 1125 UNTS 609, Art. 13(2) (hereinafter, Additional Protocol II or AP II).

[18] J. McMahan, 'The Basis of Moral Liability to Defensive Killing', (2005) 15(1) *Philosophical Issues*, 386.

[19] Arguably, there are exceptions to this general position (as when an individual fighting in an otherwise objectively unjust war uses force to justifiedly prevent a specific wrongful attack by someone otherwise fighting an objectively just war, or when a war is justified exclusively on lesser evil grounds), but they are not particularly relevant for our purposes.

[20] See Art. 4 of the Lieber Code (Instructions for the Government of Armies of the United States in the Field), General Order No. 100, Art. 14 (24 April 1863). For a book-length treatment of this principle, see L. May and J. Ohlin, *Necessity in International Law* (2014). For a more demanding account, see Y. Beer, 'Humanity Considerations Cannot Reduce War's Hazards Alone: Revitalizing the Concept of Military Necessity', (2015) 26(4) *European Journal of International Law* 801. For a leading philosophical account, see S. Lazar, 'Necessity in Self-Defense and War', (2012) 40(1) *Philosophy and Public Affairs* 3.

[21] J. McMahan, 'War Crimes and Immoral Action in War', in A. Duff et al. (eds.), *The Constitution of Criminal Law* (2013), 151, at 160.

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

Finally, revisionist just war theory has important implications vis-à-vis the *in bello* principle of proportionality. Again, proportionality in this context is generally considered to require that the anticipated harm that an act of war will cause is not excessive vis-à-vis the military advantage it will expectedly provide.[22] Namely, it prohibits an attack which may be expected to cause incidental harm to non-liable civilians which would be excessive in relation to the concrete and direct military advantage anticipated.[23] But if a given act is in pursuance of an unjust end, then it is logically impossible that it can meet the requirement of proportionality. This follows from the simple fact that there is no good to be attained by the military enterprise that can outweigh the harm the attack will cause. This is particularly so given that individuals fighting on the just side are usually not liable to being attacked. Accordingly, those who pursue unjust ends would normally fail to meet this requirement too.

To sum up, if revisionists are right, the acts of those fighting an unjust war which harm or threaten to harm those fighting on the just side (as well as any noncombatants) would typically violate the most fundamental *in bello* principles.[24] This means that revisionists radically redefine what it means to act in a morally wrongful way in war.

3. From moral wrongfulness to criminalization

Cécile Fabre, one of the foremost defenders of the revisionist position, has recently advocated a view according to which those fighting an unjust war are morally liable to be punished for murdering soldiers fighting on the just side (as well as civilians), i.e., for participating in an objectively unjustified war. Echoing Augustine, she submits that although killing or maiming those fighting an unjust war should not be deemed a criminal offence, killing those fighting on the just side should 'in just the same way as it should be a criminal offence in domestic context to kill someone without justification'.[25] Furthermore, not only political leaders and military commanders should be liable to punishment for this type of offence, as in the crime of aggression, but also rank-and-file soldiers.[26] She therefore maintains that the standard conceptualization of war crimes under international law is radically under-inclusive and that the law should be amended in order to incorporate this insight.

---

[22] See Additional Protocol I, Art. 51(5)(b). See also J. McMahan, 'Proportionate Defense', (2014) 23 *Journal of Transnational Law & Policy* 1, at 6–7.

[23] The principle of proportionality in attack is codified in Art. 51(5)(b) of AP I, and repeated in Art. 57. 'While Additional Protocol II does not contain an explicit reference to the principle of proportionality in attack, it has been argued that it is inherent in the principle of humanity which was explicitly made applicable to the Protocol in its preamble and that, as a result, the principle of proportionality cannot be ignored in the application of the Protocol.' J-M. Henckaerts and L. Doswald-Beck, *Customary International Humanitarian Law* (2005), Vol I: Rules, at 48.

[24] Except for the one of legitimate authority, which is independent of the requirement of just cause, but which is contested by some revisionist just war theorists.

[25] Fabre, *supra* note 14, at 7.3 *in fine*.

[26] Ibid.

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

Admittedly, Fabre does not believe that we should simply equate individuals fighting an unjust war with war criminals. She contains this radical implication on at least one important ground, relevant for us here. She argues that liability to being attacked and liability to being punished rest on different grounds.[27] That is, defensive killing is forward looking in that the agent seeks to block an ongoing or imminent threat. Punishment, by contrast, is backward looking; it is a response to a past wrongdoing. Accordingly, an argument for criminalization requires providing a specific justification for meting out legal punishment upon an offender.

I agree. As I have argued elsewhere, there are good reasons to believe that the scope of criminalization is largely determined by the reasons we have for punishing a particular individual.[28] We may plausibly assume here, with a significant part of the contemporary scholarship on punishment, a justification for legal punishment which rests broadly on expressivist or communicative grounds.[29] Namely, the moral value of punishment rests on the value of reflecting that offenders perpetrated a grievous wrongdoing, as a way to condemn her act. This means that, unlike liability to being attacked, liability to being punished should be the product of culpable agency. It requires the defendant satisfying the relevant *mens rea*. On these grounds, a small number of those fighting unjust wars would be exempted from criminal liability for so doing. For instance, individuals who are unable to exercise autonomous agency, such as most instances of child soldiers, or of the mentally insane or senile, would not be liable to be punished.

Notwithstanding this caveat, the key question is whether those participating in armed conflict must intend to contribute to an unjust war, or whether it suffices that they know that they are so contributing, or even that they can reasonably be expected to know on the basis of the available evidence. Fabre concludes that there are good reasons to 'reject the views that intentions or actual knowledge are necessary requirements for fault and therefore for liability to punishment'.[30] In effect, the better view seems to be that the fact that soldier A would have reasonably been expected to know that the war was unjust would suffice for her being liable to punishment.[31]

Accordingly, it would suffice for criminal liability that the accused knew that the war was unjust, or that he was required to have known. In this context, the critical issue is whether rank-and-file soldiers actually satisfy these *mens rea* standards, i.e., whether they know or 'can reasonably be expected to know that they are taking part in the wrongful killing of' those fighting an objectively just war 'when they are told, for example, to invade another country'.[32] Whether they in fact know is a matter of evidence that should be determined in trial. I will not pursue this line of argument here. The more difficult question is when we can assert that they are

[27] Ibid., at 7.2.2.
[28] See A. Chehtman, *The Philosophical Foundations of Extraterritorial Punishment* (2010), Ch. 3.
[29] Fabre, *supra* note 14, at 7.2.2. I have advocated a similar justification in note 28, *supra*, in Chapter 2. For seminal articulations, see J. Feinberg, 'The expressive function of Punishment', in J. Feinberg, *Doing & Deserving: Essays in the Theory of Responsibility* (1974); A. Duff, *Punishment, Communication, and Community* (2001).
[30] Fabre, *supra* note 14, at 7.4.1.
[31] Ibid.
[32] Ibid.

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

morally required to know (even if, as a matter of fact, they do not know) that the war is unjust. It is here that I part ways with Fabre's revisionist position.

On this point she introduces a two-pronged analysis. She recognizes that it would be unlikely that individuals living under oppressive regimes where neither ordinary soldiers nor ordinary civilians have access to open media can be required to know that the war is unjust, yet 'this does not hold for broadly democratic regimes'.[33] '[I]ndividuals under these regimes can reasonably be expected to do as much as they can to inform themselves of the justness of the war (or phase of war) which they are asked to fight and support.'[34] Furthermore, in their case, she rejects that liability would be undercut by the defence of 'superior orders' or by the so-called 'epistemic limitations objection', which claims that, given their limited access to information, rank-and-file soldiers and ordinary citizens contributing to the war effort cannot reasonably be expected to know that they are taking part in an unjust war.[35]

This position leads to a radically revisionist position vis-à-vis the laws on war crimes, insofar as it makes a large number of those fighting an objectively unjust war liable to being punished. I believe, by contrast, that adopting a revisionist position need not lead to such radical implications vis-à-vis the criminalization of wrongful conduct in war. Taking seriously a plausible *mens rea* standard requiring culpability, or at least fault,[36] would lead to a far more restrictive scope of the criminalization of wrongful conduct in war. The reason for this is that relatively few of those fighting an objectively unjust war can be reasonably expected to know, or better, are morally required to know that a war was unjust. This proposition is based on a theoretical and an empirical point.

From a theoretical perspective, I suggest that Fabre's flat rejection of the 'superior orders' defence is mistaken.[37] Although she explicitly avoids justifying this position, she stresses that soldiers, or fighters more broadly, 'are not automatons who are exonerated from the burdens of autonomous moral agency once they put on their uniform'.[38] Yet I believe we need to take into consideration the type (or basis) of authority involved, and the type of order it has issued. In short, if the order is to execute 1,000 innocent children, then the defence of superior orders would be clearly inappropriate. But an order to kill a harmless individual could provide a plausible reason for action under this type of circumstance. Consider:

> *Order:* Police officer A is ordered to detain X. While pursuing this task A is informed by his superior, who has perfect vision over the situation, that X is armed and that she is about to kill Z. He orders A to shoot at X. A shoots and kills X. X was not armed and was not threatening Z in any way. A was misguided by his superior.

It seems clear that in this case A should not be punished for his act. His action should not be censured in any way, and the reason for this is arguably that he

---

33 Ibid., at 7.4.2.
34 Ibid.
35 See Walzer, *supra* note 2, at 312.
36 See, e.g., A. Ashworth, *Principles of Criminal Law* (2009), Ch. 5.
37 I believe this is true both as a matter of normative theory and as a matter of law. I deal here with the former issue, and with superior orders in international law in Section 5, *infra*.
38 Fabre, *supra* note 14, at 7.4.2.

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

acted non-culpably. This example vividly illustrates that there are situations in which 'superior orders' are both compatible with a rational exercise of someone's autonomy and would exonerate those fighting an objectively unjust war from moral and criminal responsibility.[39] Indeed, Vitoria acknowledged that, even in situations where one party had just cause and the other did not, it could be the case that both sides are acting in good faith such that they would both be free of blame.[40]

Arguably, intuitions may not suffice to settle this issue; it seems sensible to require some kind of argument for this conclusion. In line with this example, several just war theorists have argued that soldiers who operate under a legitimate (usually democratic) authority are under a duty to fight even in an objectively unjust war. There are different justifications of this so-called 'duty of citizenship', but the upshot of this position is that fighting in such a war is, all things considered, the right thing to do.[41] Given the significance of *mens rea* for criminal liability, I suggest that for our more limited purposes of establishing criminal liability, this type of argument is at its strongest. Indeed, in a significant number of cases, participants in armed conflicts, unlike typical murderers, operate under specific institutional constraints. These settings can have a significant bearing on how we should assess their conduct.

Joseph Raz, for instance, has influentially argued that the reason why individuals should comply with an authoritative directive is that it allows them to comply better with reasons that already, objectively, apply to them if they follow the directives than if they try to follow those reasons by themselves.[42] This normative thesis is connected with the main theoretical implication of recognizing someone as a legitimate authority; namely, it provides those who are subject to it with content-independent reasons for action that are not merely to be added to other reasons but which have some kind of special, protected status.[43] This means that authoritative directives can be binding even if they are ultimately mistaken.[44] The decision to go to war usually entails a highly complex judgement which depends on an important amount of often-confidential information, as well as on complex factual and normative assessments. It is the kind of decision – we would agree – that is better left to a centralized authority. Accordingly, it follows that rank-and-file soldiers are *prima facie* bound by

[39] Although I exemplify this position by reference to a reasonable mistake of fact caused by the intervention of a legitimate authority, the same argument would apply, under certain circumstances, to mistakes of law, although this would be rarer. This has been recently discussed by the Israeli Supreme Court in *HCJ 1971/15 Al-Masri v. The Chief Military Advocate General*. Yet, this issue is beyond the scope of this article.

[40] F. de Vitoria, 'On the American Indians', in A. Pagden and J. Lawrance (eds.), *Political Writings* (1991) 231, at 282–3. Grotius reached a similar conclusion, holding that situations of invincible doubt sufficed to place both belligerents in a somewhat symmetrical position. See Neff, *supra* note 16, at 99.

[41] For different versions of this argument, see D. Estlund, 'On Following Orders in an Unjust War', (2007) 15(2) *The Journal of Political Philosophy* 213; Ryan, *supra* note 15; and most recently M. Renzo, 'Duties of Citizenship and Just War' (unpublished typescript on file with the author).

[42] See J. Raz, *The Morality of Freedom* (1988), 53; J Raz, 'The Problem of Authority', (2006) 90 *Minnesota Law Review* 1003, at 1014. I use Raz for simplicity here. My argument would also work, albeit to different degrees, with many of the other contemporary accounts of legitimate authority available in the literature.

[43] Raz argues that authoritative directives pre-empt individual decision-making; Raz, *The Morality of Freedom*, *supra* note 42, at 46. I need not take such a demanding notion of authoritative directive. I believe that Raz's conceptual understanding of authority is compatible with a slightly less rigid notion of protected reasons, such as *prima facie* or *pro tanto* reasons for action. For a recent defence of this view see Renzo, *supra* note 41, App.

[44] Raz, *The Morality of Freedom*, *supra* note 42, at 47–8.

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

a decision of a legitimate authority even if that decision is ultimately mistaken (i.e., the war is unjust).

There are three caveats to this conclusion. First, it would not apply to illegitimate authorities, i.e., those which individuals have no moral reason to obey. Despotic regimes or oppressive non-state armed groups will characteristically not be covered by this type of argument. Second, this duty of citizenship does not apply to specific individuals who *know* that the war is unjust and yet willingly participate in it. As the situation in the *Order* case illustrates, this argument would not generally apply to leaders or certain high-ranking officials. Finally, most, if not all, defenders of this duty of citizenship agree that sometimes it is both permissible and even mandatory to disobey a legitimate authority's order to go to war. The duty of citizenship is only a *prima facie* duty and it can be overridden under certain circumstances, such as when the war is 'an aberration',[45] or 'utterly irrational or even disingenuous',[46] or at least when those fighting are sufficiently confident that the war is unjust.[47]

The service conception (and almost any account of legitimate authority based on broadly epistemic or instrumental grounds) would provide us with a plausible way to accommodate this intuition. Raz acknowledges that protected reasons can be defeated, i.e., outweighed, cancelled, and even excluded.[48] In this context, we may distinguish big mistakes (binding) from clear ones (not binding).[49] In the case of clear mistakes, it would be mandatory for individual soldiers to disobey.

In short, then, in cases where the war is not manifestly unjust, individual soldiers would be under a *pro tanto* duty to comply with the order to fight, particularly those fighting for legitimate regimes. This duty shows that even if they are fighting an objectively unjustified war, they would not satisfy the *mens rea* requirement, as it would be implausible to suggest that they were morally required to have known that the war was unjust. In fact, their epistemic position cannot be easily distinguished from that of those fighting an objectively just war, since under most circumstances, and unless the war is to be considered manifestly just, they too would be relying on the judgment of the relevant authorities. They – each of them – can very seldom be absolutely certain that they are fighting with just cause and that the war is necessary and proportionate.

This argument is further expanded by a second, wider, empirical point. Namely, we have strong reasons to take seriously the so-called 'epistemic objection', i.e., the claim that ordinary, rank-and-file soldiers and citizens cannot reasonably be expected to know that they are taking part in the wrongful killing of innocent enemy (just) soldiers. Even if we accept that individuals under broadly democratic regimes are expected to do as much as they can to inform themselves on the justness of the war, the obstacles for rank-and-file soldiers to acquire the relevant empirical information regarding the justness of the war are far more serious than is commonly assumed. Although one would expect this type of limitation to be particularly severe in

[45] Ryan, *supra* note 15, at 30, 32.
[46] Estlund, *supra* note 41, at 232.
[47] Renzo, *supra* note 41, at 17.
[48] J. Raz, *Practical Reasons and Norms* (1999), 47.
[49] Raz, *The Morality of Freedom*, *supra* note 42, at 62.

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

non-democratic, oppressive regimes, I suggest it is also very significant in open, liberal democracies. The reason for this does not stand, as it is often assumed, merely on the grounds that much of the empirical information relevant for the issue of whether war would be just is classified, or involves complex calculations. Rather, the reason for this is much simpler, but also far more pervasive.

Recent empirical research has shown that our cognitive abilities, i.e., our perception of the relevant facts are generally biased by things that are beyond our control even in contexts in which people have access to information.[50] Joshua Greene argues that this is, to a large extent, the result of perception biases. One type of relevant bias is simple self-serving bias. Namely, when the facts are ambiguous, people tend to favour the interpretation of the facts that is more in line with their interests. Sports fans often experience this type of reaction vis-à-vis controversial decisions by umpires or referees. Similarly, there is a type of bias that makes us far more aware of the pain or harm we suffer as a result of someone else's acts than the pain that others suffer by our actions. This tends to result in the escalation of conflicts.[51] But critically for our purposes, people's perceptions of the relevant facts are also deeply biased by the political, social, and moral commitments they have as individuals, or by the 'tribal' groups to which they belong. This process is reinforced by informational and reputational influences which affect both factual judgements and political views, and can lead to group polarization and even to conspiracy theories.[52]

Greene illustrates these phenomena through an example that is particularly apposite for our purposes. Namely, he argues that the fact that most critics of the US invasion of Iraq in 2003 could not understand why so many Americans supported the war was the result of them not fully appreciating that 'a majority of Americans at the time believed that Saddam Hussein had been personally involved with the [9/11] attacks'.[53] Similarly, he adds, a majority of people in Jordan, Egypt, and the Palestinian territories believed that someone other than al-Qaeda (typically the US or Israeli government) had been behind the attacks. This type of phenomenon can also be seen in several other domains, such as climate change or mechanisms for the disposal of nuclear waste, among others.[54] Factual propositions which might be clearly mistaken for experts may be deeply intuitive for the lay person. In fact, Greene argues, once false beliefs become 'culturally entrenched … [they] are very difficult to change, and changing them is no longer simply a matter of educating [or providing information to] people'.[55] Accordingly, having access to a free press and a robust public arena would hardly suffice for individuals to be aware of the relevant non-moral facts in this type of politically charged scenario. It is therefore often unlikely that rank-and-file soldiers (i.e., young nationals of the relevant states, usually from socio-economic underprivileged backgrounds) are actually able to

[50] J. Greene, *Moral Tribes: Emotion, Reason and the Gap Between Us and Them* (2013), Ch. 3.
[51] Ibid.
[52] C. Sunstein, 'The Law of Group Polarization', (2002) 10(2) *The Journal of Political Philosophy* 175.
[53] Greene, *supra* note 50. See also C. Sunstein and A. Vermeule, 'Conspiracy Theories: Causes and Cures', (2009) 17(2) *The Journal of Political Philosophy* 202.
[54] Greene, *supra* note 50.
[55] Ibid.

determine the unjustness of the war with the necessary degree of certainty to put them under an obligation not to fight.

On the basis of the two arguments put forward in this Section, it hardly seems to follow that, even from a revisionist stance, individuals fighting an objectively unjust war would be *generally* liable to be punished for so doing. The reason for this is that they would generally lack the relevant *mens rea*. By contrast, I suggest that only those fighting a manifestly or patently unjust or abhorrent war would be morally liable to being punished. Interestingly, Augustine, Vitoria, and Suárez already defended a similar standard when they argued that soldiers are expected to fight unless 'the cause in which they were enlisted was patently unjust'.[56] Vitoria even added that if the war is manifestly unjust:

> the subject … must not fight, even if he is ordered to do so by the prince. This is obvious, since one may not lawfully kill an innocent man on any authority, and in the case we are speaking of the enemy must be innocent.[57]

In this context, the fact that any given war is fought manifestly without just cause, or through clearly disproportionate means, as well as through systematic rape, or through intentionally targeting unarmed civilians, would suffice to require those who fight in it to be aware that they are perpetrating a grave moral wrong by merely participating in it. To illustrate, a soldier fighting in Poland for the Nazis or in Hungary for the Soviets, a cadre fighting for ISIS in Syria or Iraq, or a rebel fighting for a bloody warlord such as Joseph Kony in Uganda or Jean Pierre Bemba in the Central African Republic would all be liable to be punished. I suggest that in each of these situations, taking part in these forces executing as many innocent civilians and openly violating so many fundamental human rights must not only be considered morally wrong, but also the type of conduct that is appropriately criminalized.[58] This would obtain irrespective of the existence of an open press or any other condition for a robust public deliberation.[59]

4. *IN BELLO* MORALITY VERSUS THE LAWS ON WAR CRIMES

Even within the revisionist camp, some would resist my previous conclusion. Jeff McMahan, for instance, believes that there are countervailing considerations that advise against criminalizing participation in an objectively unjust war.[60] By contrast,

---

[56] Neff, *supra* note 16, at 57 (and its references in footnote 62).

[57] F. de Vitoria, 'On the law of war', in A. Pagden and J. Lawrence (eds.), *Political Writings* (1991), 293 at 307, cited in Estlund, *supra* note 41, at 214.

[58] As I indicated in Section 2, *supra*, that such war systematically violates *in bello* rules suffices to turn it manifestly unjust or morally abhorrent.

[59] It may be argued that this is compatible with their moral responsibility being mitigated in certain cases by situations of extreme duress. Yet this qualification would imply that they are liable to being punished, and would affect the extent of that liability. For a good discussion of the difference between justification and total excuse, see E.R. López, 'Can there be Full Excuses for Morally Wrong Actions?', (2006) 73(1) *Philosophy and Phenomenological Research* 124.

[60] It is interesting to point out that at the level of the deep morality of war he is sympathetic to the claim that acts of almost every type of participant in an objectively unjustified war are akin to war crimes. See McMahan, *supra* note 21. The problem is, as I have suggested, that such revisionist position would be implausibly over-inclusive.

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

he advocates a neutral or symmetrical *law* on war crimes that criminalizes 'a form of action in war only when doing so would have the expected effect of reducing the amount of wrongful harm inflicted in war or, equivalently, the sum of weighted rights violations'.[61] Put differently, although he accepts that such conduct would be morally wrongful, McMahan argues that there are prudential, practical and moral reasons not to criminalize mere participation in an unjust war. He puts forward six cumulative objections which I shall consider and reject, at least insofar as they seek to undermine the specific position I advocate in this article.

First, McMahan argues that legal punishment would be unwarranted against those fighting an unjust war given the fact that individual soldiers are often neither well enough informed nor otherwise qualified to determine that the war is unjust. In many cases – he adds – 'their restricted epistemic situation is an excusing condition that is sufficient to exempt them from liability to punishment'.[62] As it is clear from my argument in the previous Section, I agree with the general thrust of this first objection. Nevertheless, it hardly leads to the purported conclusion. The fact that several, and even most, of those fighting an objectively unjust war are not in a position to conclude that the war is unjust hardly means that all of them are. Consider, for instance, militiamen fighting for Jean-Pierre Bemba in the Democratic Republic of the Congo and systematically using rape as a tactic of war. Or consider those fighting for ISIS, throwing gay people out of buildings or ordering children to carry out executions in cold blood. Consider further a Nazi soldier in the midst of occupied Poland. Even those individuals who were not directly involved in massacres, rapes, or atrocities more generally would be unable to claim that they were neither well informed nor qualified to determine that such wars were unjust (and their contribution to that enterprise morally wrong). The same would apply when a particular state or group acts manifestly without just cause.[63]

Moreover, even if any of these fighters were convinced that the war was just (and that such tactics were justified), something that in some cases we can empirically discard,[64] that kind of judgment would be entirely unreasonable. Accordingly, just as in domestic settings we hardly question the state's right to punish a serial killer who has a sincere, albeit unreasonable, belief that he is following God's commands or the orders of a particular cult, we should similarly consider this type of offender liable to being punished (provided she is not mentally insane).

Second, it may be argued against the position hereby advocated that those fighting an objectively unjust war usually act under duress: if they refuse to fight, they are

---

[61] McMahan, *supra* note 21, at 177.

[62] Ibid., at 172.

[63] Admittedly, it is difficult to find historical examples of belligerents fighting a manifestly unjust war exclusively which did not also systematically violate *in bello* rules. But this is only because belligerents manifestly violating *ad bellum* principles also tend to systematically violate *in bello* rules. Soviet military action in Hungary in 1956 may be a plausible example of the former. See C. Gati, *Failed Illusions: Moscow, Washington, Budapest, and the 1956 Hungarian Revolt* (2006).

[64] For example, Baez et al., argue that in the Colombian conflict a majority of former fighters 'joined paramilitary groups for economic reasons', and only 13 per cent had an ideological motivation for joining. See S. Baez et al., 'Outcome-oriented moral evaluation in terrorists', (2017) 1 *Nature Human Behavior* 4. Yet, see also Á. Gómez et al., 'The devoted actor's will to fight and the spiritual dimension of human conflict', *Nature Human Behavior*, available at www.nature.com/articles/s41562-017-0193-3.pdf.

threatened with punishment.[65] As a result, a threat from an external source to punish them is less likely to deter them from fighting. One problem with this objection is that it assumes that deterrence is the only, or at least the main, grounds on which liability to legal punishment must be justified. This is a controversial stance to take, and one which I reject as a matter of principle.[66] However, the main problem with this objection is that, even in its own terms, it is either unfounded or it simply proves too much. That is, whether international criminal law can deter wrongful behavior in war is an empirical question. Recent empirical research has shown that legal prohibitions have some deterrent effect in relation to wrongful conduct in war, at least in cases of mass atrocity crimes.[67] Insofar as we may plausibly assume that duress is present also in contexts of mass atrocity,[68] it follows that legal prohibitions would be able to deter participation in a manifestly unjust war. In this respect, the objection seems to stand on dubious empirical assumptions.

If, by contrast, we take this argument at face value, it fails also because it would entail decriminalizing orthodox war crimes, such as mass rape, taking of hostages, or widespread torture when performed by individuals within atrocious regimes. Insofar as these acts are also perpetrated by groups and in contexts that exercise enormous psychological pressure on soldiers, and insofar as the consequences of not participating in, or contributing to, them would be dire to these individual soldiers, this objection would be committed to not considering this type of conduct as war crimes either. This implication is something McMahan would certainly want to resist.

The more persuasive objection is that criminalizing participation in an unjust war would create incentives for prolonging that war.[69] Namely, the threat of punishment may deter individuals from surrendering. This assumption is corroborated in the empirical literature.[70] However, even if we concede this premise, it hardly warrants the proposition that mere participation in a manifestly unjust war should *never* be criminalized. Ultimately, this objection seems to assume an implausible version of the revisionist approach, i.e., one which claims that being liable to being punished entails that it is *always* justified to punish someone all-things-considered. Yet this conclusion need not follow.

Liability to being punished is compatible with reduced and alternative sanctions, but also with amnesties. Put differently, this objection can be easily accommodated by acknowledging that any plausible account of *jus post bellum* must allow for trade-offs between punishment and other goals, such as peace, truth, and reconciliation. This is the dominant position in the so-called transitional justice literature, and it applies also to the orthodox conception of war crimes.[71] Accordingly, Article 6(5)

---

[65] McMahan, *supra* note 21.
[66] See Section 3, *supra*, and Chehtman, *supra* note 28, at 43–6.
[67] See H. Jo and B. Simmons, 'Can the International Criminal Court Deter Atrocity?', (2016) 70(3) *International Organization* 443.
[68] E.g., *Prosecutor v. Erdemovic*, Judgement, Case No. IT-96-22-A, A. Ch., 7 October 1997.
[69] McMahan, *supra* note 21, at 173.
[70] See, e.g., D. Reiter and A.C. Stam, *Democracies at War* (2002), 65–9.
[71] See, e.g., J. Elster, *Closing the Books: Transitional Justice in Historical Perspective* (2004).

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

of Additional Protocol II encourages authorities to 'endeavor to grant the broadest possible amnesty to persons who have participated in the armed conflict'.[72] Yet adopting this type of framework would conceptually pre-suppose soldiers are liable to being punished for participating in a manifestly unjust war, and is compatible with calling at least some of them to account.

Fourth, McMahan objects that until international legal institutions are able to provide those participating in armed conflict with authoritative guidance on whether their side is or would be legal or illegal before or during the course of a war, it would be unfair to hold them liable to legal punishment.[73] However, this argument does not prove enough. The fact that we currently lack such legitimate authority and that in all probability there will not be one for the foreseeable future does not mean that it will always be unfair to hold certain participants in an armed conflict liable for fighting in an objectively unjust war, particularly when this was manifestly so. To repeat, those fighting for ISIS, Nazi Germany, or for the Lord's Resistance Army, should be as much liable for so doing as those responsible for the Rwandan genocide or Chilean torture, i.e., for a collective action which was, by any plausible account, morally abhorrent.

Fifth, McMahan claims that as long as international institutions remain inadequate, punishment involves the risk of victor's justice, i.e., the risk that those fighting a just war are punished by a victorious, unjust adversary.[74] This proposition seems unpersuasive. For one, I believe it is an overstatement to claim that victorious belligerents have great incentives to conduct criminal trials against the vanquished. Criminal trials are costly and they standardly provide a forum for defendants to air their grievances and defend their positions.[75] It is no coincidence that a powerful country like the US has put to trial only a handful of defendants in the multiple conflicts it identifies as its 'war against terror'. Similarly, Churchill's preference for summary executions over trials on grounds of expediency is also well-known. This seems to indicate that even if the risk McMahan identifies exists, it is largely overstated. But more importantly, this concern is not peculiar to this type of prosecution but rather relevant to war crimes trials more broadly.[76] That is, insofar as it would entail eliminating criminal responsibility for traditional war crimes, this objection simply proves too much.

Finally, it may be objected that no state could be expected to surrender a large number of its citizens for trial for doing what it had commanded them to do.[77] It would have to be coerced, and it would be absurd to suppose that compelling a recalcitrant state to extradite its former fighters for war crimes trials could be a just cause for a further war. Again this objection assumes an implausibly broad version of the revisionist position according to which every participant in an objectively

[72] Additional Protocol II.
[73] McMahan, *supra* note 21, at 173.
[74] Ibid.
[75] See, e.g., M. Koskenniemi, 'Between Impunity and Show Trials', (2002) 6 *Max Planck Yearbook of United Nations Law* 1.
[76] See, e.g., G. Simpson, *Law, War & Crime* (2007), 16.
[77] McMahan, *supra* note 21, at 173.

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

unjust war must be punished. Yet I only advocate criminalizing the conduct of only a small fraction of those actually participating in objectively unjust wars. Namely, only those participating in a knowingly or manifestly unjust war would be liable to be punished. Furthermore, the fact that they are liable to being punished does not mean that they should be punished all-things-considered. Whether and which of them should be prosecuted will certainly depend on the political and operational feasibility of conducting fair trials, as well as on the overall costs of pursuing this type of *jus post bellum* policy.

In sum, there are two related problems with McMahan's critique. First, he identifies the revisionist position as one for which *every* individual fighting in an objectively unjust war would be liable to being punished. I have argued that such a position would be untenable, and it might be unpersuasive partly for the reasons he suggests. However, from the plausible claim that it would be morally and prudentially wrong to hold accountable every participant in an unjust war, he seems to jump to the equally implausible view that none of them should be held accountable. Second, an underlying problem with these objections is that they seem to assume a situation of war as exclusively between states with similar capabilities and relative standing. However, wars need not be, and in fact increasingly are not, of this sort. Wars can be waged between state belligerents with massive differences, but also between them and different sorts of non-state armed groups (or between these groups alone). They can be waged between organizations with extremely diverging claims to legitimacy, or to legitimate authority over those who fight, and through very different tactical and strategic approaches. Once we take into consideration these further possibilities it seems not only morally sound but also practically plausible to criminalize certain types of participation in certain types of unjust wars. Perhaps the clearest illustration of the plausibility of the claim I advocate is that international law admits that under certain circumstances mere participation in an unjust war may warrant criminal liability. I will articulate this view in the following Section.

5. The crime of participating in a manifestly unlawful war: From morality to law

I have argued that the most compelling articulation of the revisionist position in just war theory would require criminalizing participation in a manifestly unjust or morally abhorrent war, or in a war a participant knows is unjust. In legal terms, the position I advocate is that individuals who participate in a manifestly unlawful war *should* be liable to be punished.[78] Of course more work is needed on how to define such a crime: both 'participate' and 'manifestly' would need to be much more precisely defined. Certain underlying offences would need to be identified (killing, maiming, and so on). Yet my aim here is not to advocate for a legal definition of a

[78] I assume for present purposes a rough overlap between manifestly unjust and manifestly unlawful wars. This is an important and sensitive assumption, which is beyond the scope of this article.

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

new war crime, but rather to provide a rationale for why this type of behavior should be rightly considered criminal under international law.

I believe this position is not only morally sound but practically and prudentially sensible, with the following caveats. First, the fact that we consider participants in manifestly unjust wars liable to being punished does not mean that they should always be punished all-things-considered. Second, it also does not mean that punishment should always be conceived as lengthy prison sentences, or that each one of them should be punished. Third, this suggestion is compatible with acknowledging that their liability can be mitigated by other considerations, most notably duress.[79] But all these caveats in fact presuppose that individuals who participate in a manifestly unjust war should be nonetheless liable to being punished; that is, that they would not be wronged if punished.[80]

This Section shows that the account I favour is much more compatible with the existing legal framework than it is assumed by *both* revisionist and orthodox just war theorists. For one, the specific standard of manifest illegality has been explicitly adopted for similar purposes under international law. For example, the provision on the crime of aggression, which was adopted as a proposed amendment to the Rome Statute, provides that the act of aggression must be a 'manifest' violation of the UN Charter.[81]

Similarly, the defence of superior orders under the Rome Statute requires *inter alia* that the 'order was not manifestly unlawful'.[82] I believe it is a mistake to assume that the defence of superior orders is simply debunked as a matter of international law. Three observations may suffice to make my point. First, many courts – both national and international – have accepted that certain superior orders may exonerate an individual from criminal responsibility for war crimes, as illustrated by Article 33(1) of the Rome Statute.[83] Although in Nuremberg both the Statute of the International Military Tribunal (IMT) and its jurisprudence rejected this defence, this is most plausibly understood as a consequence of the type of defendants that were to be prosecuted before the IMT, i.e., superior officers.[84] Second, many who

---

79 It assumes, of course, they would satisfy other conditions such as mental capacity. Indeed, this argument does not necessarily include child soldiers or other analogous cases.

80 Whether they should be punished or not and what kind of penalties should be imposed on them largely depends on the specifics of the situation. I am happy to concede that other considerations would be relevant to determine what to do, such as the prospects of a sustainable peace, the needs for reconstruction, and so on, and that in a significant number of cases it would be counterproductive to punish them.

81 Art. 8*bis* (1), Res. RC/Res. 6, The crime of aggression, Annex I, adopted on 11 June 2010. This definition, specifically the requirement that the violation of *ad bellum* rules be 'manifest', remains controversial.

82 Rome Statute, Art. 33(1)(c).

83 See references in Y. Dinstein, *The Defence of 'Obedience to Superior Orders' in International Law* (1965) and, more recently, M. Osiel, *Obeying Orders: Atrocities, Military Discipline and the Law of War* (1999).

84 Gaeta highlights that the British, French and Russian prosecutors in Nuremberg argued that this defence was not possible due to the *manifest illegality* of the orders, rather than simply relying on Article 8 of the London Charter. P. Gaeta, 'The Defence of Superior Orders: The Statute of the International Criminal Court versus Customary International Law', (1999) 10 *European Journal of International Law* 172, at 180 (emphasis added). Similarly, Garraway argues that the drafting history of the Nuremberg Charter at several points included a provision on superior orders, and suggests, quoting General Nikitchenko, that the reason they were not considered a valid defence in that context was that those particular defendants could not be carrying out orders of a superior. C. Garraway, 'Superior orders and the International Criminal Court: Justice delivered or justice denied', (1990) 81(836) *International Review of the Red Cross* 785, at 786. Put differently, Nuremberg was

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

argue that superior orders cannot exonerate responsibility for war crimes under international law do so on the implausible assumption that they are always manifestly unlawful.[85] Yet, as illustrated in Section 2, this assumption is mistaken. Finally, insofar as the existence of a legitimate authoritative directive can conceptually have an impact on the reasonable expectation that a particular individual seeks more information about the justness of a particular war, or action within the war, international criminal law should exonerate certain individuals on grounds of their being under a reasonable mistake of fact, which was itself based on the existence of such an authoritative order.[86]

Admittedly, neither the criminalization of aggression nor the limitations to the defence of superior orders captures the specific conducts at stake here, at least not vis-à-vis rank-and-file soldiers and civilians who directly take part in a knowingly or manifestly unlawful war. The former is restricted to persons 'in a position effectively to exercise control over or to direct the political or military action of a State',[87] while the latter applies to the orthodox idea of war crimes, as limited to particularly grave violations to the *jus in bello*. Nevertheless, I will now suggest the laws of armed conflict allow for the criminalization of rank-and-file soldiers for the underlying acts that constitute *merely* participating in a knowingly or manifestly unlawful war.

The main divide in the regulation of armed conflicts under international law is between international and non-international armed conflicts.[88] In international armed conflicts combatants on all sides have immunity against being prosecuted for participating in the war; they are considered privileged, and are not liable to being punished (as prisoners of war) as long as they do not breach the specific rules of a 'fair' fight.[89] This includes the regular armed forces of the state as well as other non-state combatants fighting as *de facto* organs of the state or acting under its effective control.[90] By contrast, individuals fighting for non-state armed groups in a conflict

---

in some way a *rejection* of the traditional position on the defence of superior orders which, as Oppenheim wrote in the first edition of his book, stated that '[i]n case members of forces commit violations ordered by their commanders, the members may not be punished, for the commanders are alone responsible, and the latter may, therefore, be punished as war criminals on their capture by the enemy'. Cited in Garraway, at 786. Consistent with this, in the *High Command* case (11 NMT, at 506) the judges acquitted the defendants on the basis that orders were not manifestly unlawful or known to be unlawful (even though Control Council Law No. 10 followed the IMT Statute on the rejection of superior orders).

85 Gaeta, *supra* note 84, at 191.

86 Walzer recognizes that the defence of superior orders can be based on ignorance. Walzer, *supra*, note 2, at 312. Similarly, Dinstein argues that the superior orders which are not manifestly unlawful affect individual awareness of the illegality of an act. Dinstein, *supra*, note 83, at 27–8. More in line with the proposal defended here, Cryer suggests that 'the manifest illegality test is a way of determining if the defendant ought to have known that the order was illegal'. R. Cryer, 'Superior Scholarship on Superior Orders', 2011 *Journal of International Criminal Justice* 959, at 963.

87 Rome Statute, Art. 33(1)(c).

88 See, e.g., the contributions to E. Wilmshurst (ed.), *International Law and the Classification of Conflicts* (2012), especially those in Part II.

89 Although they can be detained for the duration of the hostilities, this detention is exclusively to prevent them from further participating in the armed conflict. It cannot and should not be construed as a criminal sanction. For further details, see Geneva Convention III Relative to the Treatment of Prisoners of War, 75 UNTS 135, 12 August 1949 (hereinafter, Geneva Convention III).

90 See Articles on the Responsibility of States for Internationally Wrongful Acts, adopted by the International Law Commission on 10 August 2001, *Report of the International Law Commission, Fifty-third Session*, UN Doc. A/56/10, Arts. 4 and 8, respectively.

not of an international character are typically not protected from being punished by the belligerent state for 'merely' participating in the conflict under its own domestic law. They are liable to be punished for the perpetration of the underlying acts that constitute their participation in the armed conflict.[91]

There are, however, two (controversial) qualifications to this clear-cut position. First, under Additional Protocol I to the Geneva Conventions, the privilege enjoyed by state combatants in international armed conflicts was extended to members of certain non-state armed groups, namely, those fighting 'against colonial domination, alien occupation or racist regimes'.[92] Although this particular provision has been criticized and actively resisted by certain states (and, in fact, it may have never been successfully invoked) it has been ratified by the majority of states in the international community.[93] Arguably, the underlying reason for their protection against prosecution was precisely that their fight was increasingly considered lawful, or at least justified. The *Commentary to API* by the International Committee of the Red Cross states:

> [T]he struggle of [these] peoples … is legitimate [and] any attempt to suppress such a struggle is incompatible with the Charter, the friendly Relations Declaration, the Universal Declaration of Human Rights, and the Declaration on the Granting of Independence, and constitutes a threat to international peace and security.[94]

Accordingly, these non-state fighters can also be considered privileged combatants under the laws of armed conflict, and therefore are not liable to being punished for the underlying acts that constitute merely participating in the armed conflict.

Second, some international law scholars and some courts have argued that by abusing the laws of armed conflict certain combatants become liable to prosecution.[95] I refer both to individuals who have forfeited their right to combatant status or to the status of prisoner of war, as well as to those civilians who have lost or

---

[91] International law does not, itself, forbid members of non-international armed groups to fight. See text corresponding to note 104, *infra*.

[92] AP I, Art. 1(4). The way in which the law has rendered this outcome is by claiming that conflicts of this particular type are 'international armed conflicts'.

[93] This was one of the reasons the US declined ratifying Additional Protocol I. See, e.g., Letter of Transmittal from Ronald Reagan, President of the United States, to United States Senate (20 January 1987), reprinted in (1987) 81 *American Journal of International Law* 910. This provision has also been included in relevant reservations by several countries, including France, the UK, Belgium, the Republic of Korea, Ireland, and Canada. On its lack of effective use, see D. Akande, 'Classification of Armed Conflicts: Relevant Legal Concepts', in *supra* note 88, at 49). Yet a rational reconstruction of the law certainly allows for considering this provision binding, at least as a matter of treaty law.

[94] Y. Sandoz et al. (eds.), *Commentary to the Additional Protocols of 9 June 1977 to the Geneva Conventions of 12 August 1949* (1987), 46; R. Sloane, 'The Cost of Conflation: Preserving the Dualism of *Jus ad Bellum* and *Jus in Bello* in the Contemporary Law of War', (2009) 34 *Yale Journal of International Law* 47, at 65, and references therein. Other people invoke other grounds for this particular extension. Corn, e.g., suggests that the rationale for their different treatment was that these particular groups were generally not perceived as committing treason, and therefore did not clearly violate the sovereign right of states. G. Corn, 'Thinking the Unthinkable: Has the Time Come to Offer Combatant Immunity to Non-State Actors?', (2011) 22(1) *Stanford Law & Policy Review* 253, at 281.

[95] See Y. Dinstein, *The Conduct of Hostilities under the Law of International Armed Conflict* (2004), 29; L. Oppenheim, *International Law: Disputes, War and Neutrality* (1952), 256; G. Schwarzenberger, *International Law as Applied by International Courts and tribunals: The Law of Armed Conflict* (1968), 115–17, among others.

forfeited their civilian status, often identified as 'unprivileged belligerents'.[96] In the influential *Quirin* case of 1942, the US Supreme Court decided that unlawful German combatants could be prosecuted and punished by military tribunals for mere participation in the Second World War.[97] This classification stands on fairly firm grounds under international law, even if its precise implications are deeply contested. Under the 1899 and 1907 Hague Regulations, subsequently incorporated into Article 4 of Geneva Convention III, combatants are considered privileged only if they distinguish themselves from the civilian population (if they carry arms openly and wear a fixed distinctive symbol recognizable from a distance), operate under responsible command, and are part of an organization that respects the laws and customs of war.[98] Thus, combatants who fail to comply with these conditions would lose their shield against prosecution and become liable to being punished for participating in such a war.[99] This does not mean, however, that they are liable to being punished for violating a rule of international law, as I will discuss below.[100] But it does mean that they are liable for perpetrating any of the underlying acts which constitute participation in an unlawful war.

Accordingly, I submit that the revisionist account I advocate provides the theoretical tools for a plausible, indeed, morally sound reconstruction of existing laws of armed conflict. For one thing, this legal framework contemplates that some individuals are liable to being punished for the underlying acts that constitute participation in an unlawful war, while others are not. Although one could construe this distinction along non-state/state lines, this would not reflect the more nuanced framework available in international law. As indicated, the law regulates at least three cases of non-state armed groups as privileged combatants, i.e., not liable to being punished for participating in an armed conflict. These are cases of groups which may plausibly be described as acting *prima facie* with just cause or at least not participating in a manifestly unlawful war.[101]

Furthermore, the second qualification introduced above – that of unprivileged combatants – allows for certain *prima facie* privileged (both state and non-state)

---

[96] See, e.g., C. Garraway, 'Interoperability and the Atlantic Divide – A Bridge over Troubled Waters', 80 *International Legal Materials* 337, at 344; R. Baxter, 'So-Called "Unprivileged Belligerency": Spies, Guerrillas and Saboteurs', 1951 *British Yearbook of International Law* 323–45.

[97] *Ex parte Quirin, et al.*, 317 US 30, 31 (1942). See, similarly, the decision of the Judicial Committee of the Privy Council in *Mohammed Ali v. Public Prosecutor*, [1968] 3 All ER 488.

[98] Geneva Convention III. For a comprehensive list of the legal requirements, see Dinstein, *supra* note 95, at 37–40.

[99] Additional Protocol I, Art. 44(3). AP I not only allowed for combatants in certain non-state armed groups being considered privileged; it also controversially narrowed the category of unprivileged combatants. In short, it establishes that it sufficed for immunity from prosecution that they carry arms openly 'within each individual engagement', and they are 'made visible to the adversary while engaged in a military deployment preceding the launching of an attack' in which they would participate. Not surprisingly, this was heavily criticized by part of the international community, as well as by many prominent international law scholars. Yet, the main criticism of this extension was not that it was normatively flawed, but rather that it was 'diluting one of the most important *quid pro quos* of humanitarian law'. Corn, *supra* note 94, at 274, and references in footnote 81.

[100] I am grateful to an anonymous reviewer for pressing me on this point.

[101] It would be plausible to suggest that there may be further non-state armed groups which should be analogized to the three explicitly provided for here. Arguably, this would be the position most clearly compatible with a strong commitment with the respect and protection of fundamental human rights.

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

combatants to be prosecuted for participating in a knowingly or manifestly unlawful war. That is, insofar as not systematically violating the *jus in bello* is part of the requirements for a lawful war, as well as a condition for lawful belligerency,[102] considering these combatants liable to being punished would be entirely compatible with the particular revisionist position I am advocating.

There are, admittedly, two inconsistencies between this proposed interpretation and the normative position advocated in Section 3. On the one hand, the criteria utilized to identify an unprivileged combatant under international law do not consider a combatant fighting manifestly or knowingly in violation of the rules on the use of force (i.e., *jus ad bellum*) liable to being punished, insofar as she complies with the main *in bello* rules. By contrast, I argued that we have good reasons to consider such individuals liable to be punished at the bar of justice. Given the central role that the principle of separation between *ad bellum* and *in bello* considerations plays in the structure of incentives of the legal regulation of war, this proposal may face significant resistance.

However, this need not pose a serious problem for the account I favour. Although logically possible, it would be extremely rare for belligerents to fight in a war that is manifestly unlawful from an *ad bellum* perspective, but do so largely in accordance with *jus in bello* requirements. Manifestly unlawful wars would normally violate both sets of principles.[103] Accordingly, keeping the legal principle of separation between *ad bellum* and *in bello* considerations would not entail accepting a bar to the prosecution of a sufficiently significant number of morally liable combatants, as per my argument in Section 3. As a result, I would be prepared to accept that there would be plausible policy considerations for law not mirroring morality on this particular issue.

On the other hand, the second inconsistency concerns the legal source of this crime. Namely, unlike the regime which conceives war crimes as serious violations of the laws of armed conflict (willful killing of civilians, torture, rape, and the like), existing international law neither *directly* criminalizes participating in manifestly or knowingly unlawful wars, nor does it criminalize the underlying acts perpetrated therein; it merely removes a shield otherwise available to (lawful) combatants as a means of protection.[104] This means that, whereas a war criminal can be prosecuted under international law, an unprivileged combatant or other type of unprivileged participant in an armed conflict may only be prosecuted under the domestic law of the state. Accordingly, liability for this type of crime is subject neither to universal jurisdiction, which is one of the defining features of an international crime, nor (in principle) the jurisdiction of international criminal tribunals.[105]

---

[102] Any use of force that incurs in this type of behavior would be very unlikely to satisfy *ad bellum* necessity and proportionality. On these requirements, see, e.g., *Military and Paramilitary Activities in and against Nicaragua (Nicaragua v. United States)*, Merits, Judgment of 27 June 1986, [1996] ICJ Rep. 14, para. 176; *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion of 8 July 1996, [1996] ICJ Rep. 226, para. 41. For a more recent statement, see C. Tams, 'Necessity and Proportionality', in L. van den Herik and N. Schrijver (eds.), *Counter-Terrorism Strategies in a Fragmented International Legal Order* (2013).

[103] For a possible exception, see *supra* note 63.

[104] Dinstein, *supra* note 95, at 234.

[105] Ibid., at 237.

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

By contrast, I believe this legal regime would be a better fit for this type of offence. To illustrate, in December 2014 Spain detained and prosecuted seven individuals for recruiting women combatants and sex slaves for ISIS.[106] Belgium, in turn, has increasingly faced the challenge of jihadists returning from Syria.[107] This phenomenon is extending to many other countries, including England, Denmark, France, Germany, Italy, the Netherlands, and Australia.[108] Such individuals are currently considered liable to being punished under each of these legal systems. Some of them are monitored weekly while others are put on trial.[109] If tried and found guilty they would have to serve prison sentences. Now, these individuals are being prosecuted under local laws and for domestic offences against the laws of each of these countries.[110] Yet I find the legal basis for these prosecutions objectionable insofar as, often enough, these offences are not being perpetrated under any of the standard jurisdictional bases accepted under international law. For it could well be that defendants did not perpetrate the alleged conduct on their territory, that they are not nationals of those states, that their acts did not ultimately harm any national of these countries, nor sufficiently affect their sovereign interests.[111]

By contrast, the argument hereby favoured would entail, much more plausibly, that these types of prosecutions are warranted, but that they should be based on a violation of international law, rather than 'mere' violations of the laws of Spain, Belgium, England and Wales, etc.[112] Three further reasons support this proposition. First, killing a soldier fighting for a manifestly just cause is objectively a serious wrong. Cassese suggested this much when justifying the extension of the notion of war crimes to those perpetrated in non-international armed conflicts in the *Tadić* decision: '[A] murder is a murder, whether it is committed within the framework of an international armed conflict, a war proper, or a civil war.'[113] The extension of the

---

[106] See 'Siete detenidos por captar 12 mujeres para el Estado Islámico', *El País*, 16 December 2014, available at politica.elpais.com/politica/2014/12/16/actualidad/1418717071_972920.html.

[107] On Belgium, see cat-int.org/wp-content/uploads/2017/04/Terrorist-attacks-Report-2013-2016.pdf.

[108] On England, see, e.g., www.independent.co.uk/news/uk/politics/plan-to-charge-jihadists-with-treason-will-not-work-claims-terror-expert-9804322.html, acknowledging liability but arguing against prosecution on policy grounds. For France, see www.trt.net.tr/francais/europe/2015/05/15/france-premiere-condamnation-d-un-combattant-a-l-etranger-276470; for Denmark, see www.thelocal.dk/20160622/denmark-convicts-first-isis-foreign-fighter; for a conviction for recruitment of foreign fighters in Australia, see www.scmp.com/news/asia/australasia/article/1988886/australian-convicted-recruiting-foreign-fighters-islamic-state. On Germany, see www.spiegel.de/international/germany/germany-faceschallenges-in-putting-islamic-state-radicals-on-trial-a-991744.html; for Italy, see, e.g., eblnews.com/news/europe/italy-court-issues-first-ever-foreign-fighter-terrorism-conviction-49029; for the Netherlands, see www.bbc.com/news/world-europe-35064597.

[109] Personal communication with Belgian prosecutor (on file with author).

[110] Interestingly, the alleged charge against the UK nationals would be treason. www.independent.co.uk/news/world/middle-east/government-moots-treason-trials-for-returning-british-isis-fighters-9799968.html.

[111] In Belgium, for instance, the crime is often considered to be leaving the country with intent to participate in terrorist activity. See, e.g., decisions Tribunal de Première Instance de Bruxelles (70e), 10 March 2017; Cour d'Appel de Bruxelles (12e), 2 June 2017 (on file with author). For a critical analysis of this type of jurisdictional basis, see Chehtman, *supra* note 8, at 404.

[112] Although this article assumes that they would be largely provided within the laws on war crimes, this need not be the case. I would concede that this type of act can be subject to international criminalization on other grounds. This possibility, however, is beyond the scope of this article. I am grateful to an anonymous reviewer for pressing me on this matter.

[113] 'Nino – In His Own Words', (2011) 22(4) *European Journal of International Law* 931, at 942 (Interview with Antonio Cassese).

notion of war crimes I am proposing here is a logical corollary of that crucial step, albeit *de lege ferenda*. Second, there are practical and political considerations that would make trials before international, internationalized, or third state courts (on grounds of universal jurisdiction) often more likely to be fair in terms of due process, and more likely to convey a clear message of moral condemnation than the courts of 'interested' parties.[114] This should address, to a significant extent, McMahan's concern with victor's justice. Finally, as forcefully argued by Kathryn Sikkink, trials before international courts or third country tribunals can have a cascade effect on domestic trials, narrowing the so-called impunity gap and contributing to the international rule of law.[115]

6. Conclusion

In this article I have argued that anyone fighting a manifestly or knowingly unlawful war should be liable to be punished for any harmful consequences she brings about, or contributes to bringing about as part of her participation in the armed conflict. This means neither that she ought to be punished all-things-considered, nor that lengthy prison sentences are the only appropriate response to this kind of wrongdoing. It merely means that she would not be wronged if punishment were inflicted upon her.

This position is being advocated as the one most consistent with our underlying moral commitments, as identified by the revisionist position in just war theory. Insofar as we are ready to accept that those fighting an objectively unjust war perpetrate grave moral wrongs, it seems only natural that when they do so knowingly, or when they could not reasonably claim that they did not know that the war was unjust, they make themselves liable to the kind of moral condemnation that punishment entails. This implication seems to follow neatly from a strong commitment to basic human rights.

Furthermore, I have suggested that this position is capable of making moral sense of the existing international law to a greater degree not only than alternative revisionist approaches, but also crucially than the orthodox position in just war theory. While the former are over-inclusive in terms of the scope of criminal liability they defend, the latter seems seriously under-inclusive. Ultimately, I advocate redrawing the relevant lines in international law between those 'liable/not liable' to be punished from the mainstream 'non-state/state' criterion, to a more persuasive 'manifestly unlawful/not manifestly unlawful' one. This entails an interpretation of existing international law *de lege lata* which puts at its centre individual human beings and respect for their fundamental rights, instead of the principle

[114] At the very least, they would normally be free from 'victor's justice' and *tu quoque* accusations. On victor's justice, see, e.g., G. Bass, *Stay the Hand of Vengeance: The Politics of War Crimes Trials* (2000), Ch. 5; Simpson, *supra* note 76, Ch. 5. On *tu quoque*, see S. Yee, 'The *Tu Quoque* Argument as a Defence to International Crimes, Prosecution, or Punishment', (2004) 3 *Chinese Journal of International Law* 87.

[115] K. Sikkink, *The Justice Cascade. How Human Rights are Changing World Politics* (2011).

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press

of state sovereignty.[116] It also entails, *de lege ferenda*, that those who participate in knowingly or manifestly unlawful wars should be punished for violating international criminal law, rather than for merely violating the local laws of any particular state.

[116] See, *inter alia*, D. Luban, *Legal Modernism* (1997), 335–62; J. Alvarez, 'Nuremberg Revisited: The Tadić Case', in (1996) 7 *European Journal of International Law* 245.

https://doi.org/10.1017/S0922156517000498 Published online by Cambridge University Press