

The Incompatibility of Rawls's Justice as Fairness and His Just War Approach

VICENTE MEDINA*

Abstract. A fundamental tension exists between Rawls's ideal Kantian conception of justice as fairness (JAF), which requires respecting people as ends, and his realistic non-Kantian consequentialist conception of a supreme emergency in a just war. By justifying the targeting of objectively innocent noncombatants during a supreme emergency exception, Rawls allows for treating them as means only. Hence, his appeal to a supreme emergency is insufficient to avoid this tension. First, since for him JAF is ideal but also practical, one might argue that his fictional people in the original position must reflect on the justification for using force on behalf of JAF. And second, since Rawls justifies targeting objectively innocent people during a supreme emergency exemption, he justifies what one might conceive of as emergency terrorism. Emergency terrorism, however, treats people as means only. Therefore, Rawls's Kantian conception of JAF is in tension with his consequentialist justification of a supreme emergency in a just war and hence with emergency terrorism.

1. Introduction

I argue that there seems to be a possible and perhaps unavoidable incoherence between Rawls's Kantian conception of justice as fairness (henceforth JAF) and his realistic non-Kantian consequentialist conception of a supreme emergency during a just war. Scholars of different ideological persuasions familiar with John Rawls's oeuvre might dismiss my argument as inconsequential or at best as a purely rhetorical one. I think it is neither of the above. In my argument, I am focusing on a legitimate philosophical and political issue regarding the meaning and scope of Rawls's ideal but also practical Kantian conception of JAF as discussed in his *Theory of Justice* (henceforth TJ) and his *Political Liberalism* (henceforth PL), and its possible tension with his

* I am grateful to my colleagues David O'Connor and Travis Timmerman, who generously read drafts of this paper. Their criticisms and suggestions were most valuable. Chris Herrera, Montclair State University, also generously read drafts of this paper and made insightful suggestions for improving it. I am also grateful to my colleague Yvonne Unna, with whom I discussed issues related to Kant's categorical imperative, enabling me to have a better understanding of it. I must also thank all members of the department who supported my sabbatical application during the spring of 2021. Also, former Chair Abe Zakhem, former Dean Peter Shoemaker, and former Provost Katia Passerini deserve credit for having approved my sabbatical leave to complete this work. Our staff at the Walsh library deserve credit, too, for ordering material in a timely fashion. Finally, I must express my gratitude to the editors and reviewers of this journal for helping me improve my argument.

realistic non-Kantian version of just war as applied to real people in his *Law of Peoples* (henceforth LOP).

In TJ and PL but also in LOP, Rawls presupposes that the original position is necessary for the justification of JAF. Moreover, he assumes that his conception of justice has moral and political import. As he states: “Both *A Theory of Justice* and *Political Liberalism* try to show how a liberal society might be possible. *The Law of Peoples* hopes to say how a world Society of liberal and decent Peoples might be possible” (Rawls 1999d, 6).

Since Rawls assumes that his political philosophy has moral and political import, it is important to explore whether his views on moral and political matters, including his ideal Kantian conception of JAF and his consequentialist conception of a supreme emergency, are consistent. I offer the following argument for engaging in this project. Rawls argues for his conception of JAF as an ideal but also practical Kantian conception of justice based on his own interpretation of Kant’s formulation of the categorical imperative of humanity. In addition, he defends his own realistic consequentialist conception of a supreme emergency in his conception of just war. However, his conception of JAF requires respecting people as ends rather than as means only. Rawls argues that during a supreme emergency in a just war, it is justified to deliberately target objectively innocent people. By doing so, he seems to allow for treating people solely as means, which goes against the letter and the spirit of his conception of JAF and against Kant’s categorical imperative of humanity. Therefore, Rawls’s conception of JAF seems to be inconsistent with his consequentialist conception of a supreme emergency in a just war.

Of course, Rawls and Rawlsians could argue that a supreme emergency is such a catastrophic situation where deontological considerations need to be temporarily suspended to defeat a vicious enemy or to protect a worthwhile political community to bring about a greater good. After all, Michael Walzer, whose views on supreme emergency influenced Rawls’s view of it, has argued accordingly (Walzer 2004, 33–50). Yet the exceptional nature of the situation does not do away with the above-mentioned inconsistency. Nor does it justify choosing between the lesser of two evils because we assume that there are no more options available. In addition, the lesser of two or more evils is a still an evil, as is the case with the act of deliberate targeting objectively innocent noncombatants in war.

I divide my argument into three parts. In the first part, I underscore that Rawls does not fully acknowledge the crucial role that force, understood as political violence, might play in establishing or defending his Kantian conception of JAF embedded in actual political institutions and practices. By not fully acknowledging the above-mentioned role, Rawls’s conception of JAF might be viewed as mostly ideal and utopian rather than practical. And yet he argues otherwise.

In the second part, I explore whether Rawls’s Kantian conception of JAF could be consistent with his realistic non-Kantian consequentialist conception of supreme emergency in a just war. I argue that his conception of JAF commits him to treating people as ends. However, in his defense of a supreme emergency exemption, he allows for deliberately targeting objectively innocent people, thereby treating them as means only. Hence, it seems that his ideal Kantian conception of justice clashes with his realistic consequentialist conception of a supreme emergency. Moreover, terrorist violence has been frequently characterized as the

deliberate use of violence against objectively innocent people to try to achieve political goals. Since that is precisely the type of violence that Rawls justifies during a supreme emergency exemption, I argue that Rawls seems to justify what one might describe as “emergency terrorism.”¹

In the third part, I sum up my argument. I restate that Rawls's Kantian conception of JAF requires respecting people as ends rather than as means only. By justifying deliberately targeting objectively innocent people during a supreme emergency exception, he justifies what might be described as emergency terrorism. But the practice of emergency terrorism treats people as means only rather than as ends. Hence, there is a fundamental tension between his Kantian conception of JAF and his non-Kantian consequentialist justification of emergency terrorism. As a result, his ideal Kantian conception of JAF and his realistic non-Kantian consequentialist conception of supreme emergency within a just war paradigm seems to be mutually exclusive.

2. Rawls's Just War Approach: All Too Real for Its Idealism

In TJ, but especially in PL and in LOP, Rawls uses his conception of the original position to develop what he conceives of as a practical political conception of justice embodied in broadly construed liberal constitutional democracies. If an ideal theory of justice, like Rawls's JAF, is to have any moral import, those who defend it need to justify the moral principles embedded in it. Yet if such a theory is to have actual *political* import, those who defend it need to address not only how to establish just political institutions and practices that mimic the ideal theory but also how to actually defend them. As a last resort, one might need to justify the use of force understood as political violence to establish and defend the already mentioned institutions and practices from those who try to undermine or destroy them by engaging in morally questionable behavior, including the unjustified use of violence or the threat of it.

I grant that peaceful rather than violent means are preferable to address intrastate or interstate deep-seated grievances of individuals or groups. Still, I am assuming that peaceful means alone are sometimes insufficient to establish or defend just political institutions from domestic or international militant intolerant people or from aggressive authoritarian states.

Militant intolerant people are individuals or groups predisposed to use violence or the threat of it to challenge just institutions and/or to inflict great suffering on those who do not deserve it.² Aggressive authoritarian states are indecent regimes whose harsh internal policies violate their citizens' fundamental human rights. Also, their sometimes expansionist external policies might unjustifiably threaten the territorial integrity of other sovereign states. By doing so, they are likely to violate the fundamental human rights of citizens belonging to those sovereign states. One might conceive of aggressive authoritarian states as being equivalent to Rawls's conception of outlaw states (Rawls 1999d, 4–5 and 80–1).

¹ For different definitions of terrorism, see Medina 2015, 28–39.

² For the concept of militant intolerant people, see Medina 2010.

Rawls believes that peaceful rather than violent means ought to be the ideal norm for solving political conflicts. Hence, unlike realist political theorists such as Machiavelli and Hobbes, he assumes that political violence is an anomaly that occurs because individuals need not always act rationally or reasonably, and they might not be living within liberal constitutional democracies. Regrettably, members of militant intolerant people and aggressive authoritarian states are neither reasonable in the Rawlsian sense nor do they respect the values of such liberal democracies. They frequently use political violence as a first rather than as a last resort to undermine the liberal and democratic values of such constitutional political systems or they might even try to overthrow them.

In LOP, Rawls shifts his focus from individuals as representatives of other individuals' interests, as argued in TJ and PL, to individuals as representing the interests of well-ordered peoples. His conception of a well-ordered people is an extension of his conception of a well-ordered society as described in TJ and PL (Rawls 2001, 8–9). But, unlike his well-ordered society in TJ and PL, which applies to the domestic realm, his conception of a well-ordered people describes how such a people might establish its liberal conception of justice in the international realm within the community of nations (Rawls 1999d, 32).³

In TJ, Rawls argues that his fictional individuals are likely to choose his two liberal principles of justice, giving priority to his first principle, based on egalitarian rights, over his second principle, allowing for economic inequality only if it benefits the least advantaged members of society (Rawls 1999e, 266). Yet he admits limitations to the lexical ordering of his two principles. He writes: "In the more extreme and tangled instances of nonideal theory this priority of rules will no doubt fail; and indeed, we may be able to find no satisfactory answer at all" (ibid., 267). Still, he proposes that we "postpone the day of reckoning as long as possible" (ibid.), acting as if such a day will never come.

One can argue that with the advent of international terrorism, especially militant jihadism and the so-called war on terror, the day of reckoning has come. Others might argue that the day of reckoning has arrived, because of the threat posed not only by militant jihadism, but also by the ominous new wave of populist parties on the right and xenophobic populist leaders worldwide who, despite having been democratically elected, have used and continue to use their authoritarian disposition to try to subvert the foundations of liberal constitutional democracies. In addition, these authoritarian leaders' virulent and widespread attacks on foreigners of myriad racial, ethnic, and religious backgrounds have fomented a surge of white supremacist and neo-Nazi groups. Some might argue that I am overstating the above-mentioned threats. Still, the threat of militant jihadism as well as the threat of white supremacist and neo-Nazi groups seem to be as real as it can get. These militant intolerable people have no scruples in resorting to violence against those whom they identify as unredeemable enemies.

To try to overcome criticisms of his ideal theory in TJ, Rawls's discussion of the original position in PL and LOP refers to real individuals or groups rather than to fictional individuals. He now underscores the need to achieve an overlapping consensus among free and equal citizens. As reasonable individuals who are fairly situated, they can offer only good reasons to establish and preserve JAF in a pluralistic society encompassing the domestic and international realms (Rawls 1996,

³ Some scholars argue that Rawls's theory of justice in TJ and PL do not fit with his conception of justice in LOP. See Pogge 2006, esp. 211.

22–8).⁴ For Rawls, good reasons are public reasons—that is, a disposition by free and equal citizens to justify to one another their views of political justice that affect them as citizens rather than as private individuals, provided that all act likewise (Rawls 1999c, 213–20).

Also, in order to make the case for JAF as a practical conception, Rawls defends his just war approach in LOP. But while an appeal to reasonableness and fairness might work within his ideal theory as described in TJ, such an appeal would likely be insufficient to establish, guarantee compliance with, or defend JAF embedded in actual institutions and practices against those who are ready to challenge them by using violence—be they nonstate actors like militant intolerant people or states like authoritarian aggressive states.

In LOP, Rawls refers to international law as the law of peoples, which he calls a “*realistic utopia*” (Rawls 1999d, 4; Rawls’s italics). It is realistic because this law takes people “as they are,” and it presupposes that the principles of justice are applicable to actual political institutions, especially nation-states. It is utopian, however, because it takes “constitutional and civil laws as they might be, that, is as they would be in a reasonably just and well-ordered democratic society” (ibid., 13).

Despite his conception of a realistic utopia, Rawls defends his justification of a supreme emergency exemption during a just war. By doing so, he justifies deliberately targeting objectively innocent people during a supreme emergency. Such a practice, however, seems to violate Kant’s categorical imperative of humanity which, according to Rawls is an integral part of JAF. In addition, the practice of deliberately targeting objectively innocent people violates the principle of non-combatant immunity (PNCI) as found in traditional Christian just war theory and in contemporary just war theory as expressed in international humanitarian law (IHL).

Rawls’s moral sensitivity and respect for human dignity led him to publicly criticize what he viewed as morally questionable practices, such as the US fire-bombing of Japanese cities, especially the dropping of the atomic bombs on Hiroshima and Nagasaki on August 6th and 9th, respectively, in 1945 at the end of World War II. For him, these acts “were very great wrongs” (Rawls 1999b, 565; cf. Rawls 1999d, 95). In LOP, however, Rawls refers specifically to the just war doctrine to criticize the above-mentioned wrong acts. Rather than talk about “decent democratic nations,” he now uses the concept of “well-ordered peoples” to refer to liberal constitutional democracies and illiberal but decent societies (Rawls 1999d, 63).

For Rawls, well-ordered peoples are people who embrace his liberal conception of a constitutional democracy or those whom he labels as decent. He assumes that either liberal or decent peoples will settle their disputes peacefully and must respect human rights (ibid., 4).⁵ According to him, all liberal people are decent people but not the other way around. Hence, he also identifies the existence of decent *nonliberal* people. But he insists that all decent people, be they liberal or not, “must

⁴ For Rawls’s conception of the original position in the international realm, see Rawls 1999d, 30–4.

⁵ For Rawls’s taxonomy of liberal and decent people but also outlaw states, see Rawls 1999d, 4.

honor the laws of peace” and “its system of law must [...] respect human rights” (ibid., 67).

As recognized by Rawls himself, his conception of just war in LOP has been influenced by Michael Walzer’s conception of it (ibid., 95 n. 8). Yet their conceptions also differ in relevant respects. Walzer argues that during a supreme emergency, soldiers and statesmen can justifiably violate the PNCI to try to salvage their political community omitting any consideration about the moral qualities of the community. Rawls, however, tries to fix Walzer’s omission by justifying violating the PNCI during a supreme emergency in a just war only when it is necessary to protect the existence of well-ordered peoples (Walzer 1992, 254, 268, 326).⁶

Rawls offers six principles or conditions that well-ordered peoples might appeal to when they engage in a just war against, for example, authoritarian aggressive states or outlaw states. While he refers to these principles or conditions restricting the conduct of war as *jus in bello* (i.e., justification of policies during war), they are a mélange of ideas, including *jus ad bellum* (i.e., justification for waging war) and *jus in bello*, and Kant’s suggestions for making a future peace possible.⁷ The following are abbreviated versions of Rawls’s proposed principles or conditions as they appear in LOP:

- (i) The aim of a just war waged by a just well-ordered people is a just and lasting peace [...].
- (ii) Well-ordered peoples do not wage war against each other [...].
- (iii) In the conduct of war, well-ordered peoples must carefully distinguish three groups: the outlaw state’s leaders and officials, its soldiers, and its civilian population. [...]
- (iv) Well-ordered peoples must respect, so far as possible, the human rights of the members of the other side, both civilians and soldiers [...].⁸
- (v) [...] Although all the preceding principles also specify duties of statesmanship, this is peculiarly true of (iv) and [...] (v). The way a war is fought and the deeds done in ending it live on in the historical memory of societies and may or may not set the stage for future war. [...]
- (vi) Finally, practical means-end reasoning must always have a restricted role in judging the appropriateness of an action or policy. This mode of thought [...] must always be framed within and strictly limited by the preceding principles and assumptions [...]. The only exception is situations of *supreme emergency* [...]. (Rawls 1999d, 94–7; italics mine)

The first condition resembles a concern of classical pagan and Christian just war traditions. In addition, by accepting the hypothesis embedded in his second condition, Rawls embraces Kant’s Second Definitive Article for a Perpetual Peace. Kant assumes that once nations agree to enter into a federation, they are likely to move gradually into a league of peace, which at some point in the future will result in a state of perpetual peace (Kant 1991b, 98–108, esp. 102–4). In the spirit of Kant, Rawls assumes that well-ordered peoples do not wage war among themselves. They are rather

⁶ For an illuminating argument between the substantive differences that exist between Rawls’s and Walzer’s views of supreme emergency, see Roberts 2011.

⁷ There is a contemporary debate between those who view Kant as part of the just war tradition and those who view him as a critic of this tradition. For the first view, see Orend 2000. For the second view, see Williams 2012. It is unclear to me whether Rawls favors any of the above views.

⁸ Rawls maintained “that [civilians] can never be attacked directly except in times of *extreme crisis*” (Rawls 1999b, 567; italics mine). In his new conception of well-ordered peoples, he omits this exception. In LOP, however, he still argues that during a “supreme emergency” civilians *can* be targeted (Rawls 1999d, 98).

provoked into war by aggressive outlaw states, such as authoritarian aggressive states, whose belligerent policies threaten the security and stability of well-ordered peoples (Rawls 1999d, 51).⁹

The third condition embodies the principle of discrimination, or the PNCI, which is shared by *jus ad bellum* and *jus in bello* in the contemporary law of armed conflict (LOAC). The fourth condition resembles a core concern of the classical Christian just war tradition against the deliberate killing of objectively innocent people (Aquinas 2006, 39 [IIa-IIæ, q. 64, a. 6]; see also Vitoria 1991, 316 [q. 3.2, § 38]). But it also shows Rawls's concern with protecting human rights in war, including the rights of both combatants and civilians.

The fifth condition resembles Kant's sixth article for making a perpetual peace possible. The parties involved in a violent conflict must avoid engaging in actions that are likely to undermine trusting one another by employing morally suspect practices, like using assassins, poisoners, or engaging in breach of surrender, among other questionable practices (Kant 1991b, 96). Following Kant, Rawls presupposes that how people fight a war can have future repercussions for establishing either a truce to prepare for a future war or a long-lasting peace.

Rawls's sixth condition allows for a restricted use of practical means-end reasoning found in the principle of proportionality in LOAC and shared by both *jus ad bellum* and *jus in bello*. The untoward consequences that result from the use of force in trying to bring about a state of peace must not exceed the calamities of the unjust state of affairs that one is trying to overcome.

By adopting the above six principles or conditions, Rawls defends his own version of a just war paradigm. Still, the justification of war is always a challenge because we have reason to believe that, as a result of it, many objectively innocent people are likely to be killed or seriously harmed. In his case, the already mentioned challenge presents him with a philosophical conundrum: He holds a Kantian conception of justice—namely, JAF, based on his own interpretation of Kant's categorical imperative of humanity—and simultaneously tries to justify targeting objectively innocent noncombatants during a supreme emergency in a just war. Such a justification contradicts the letter and spirit of Kant's categorical imperative of humanity (Kant 1964, 96 [AK 4:429]). In addition, Rawls ignores Kant's harsh criticisms of the just war tradition. While Kant allows for defensive wars, he writes that “reason [...] absolutely condemns war as a test of rights and sets up peace as an immediate duty” (Kant 1991b, 102–5; cf. 1991a, 150–7 [AK 6:343–51]). Thus, whether Rawls succeeds in his justification of a supreme emergency exemption without creating an internal tension within his own Kantian theory of justice seems doubtful.

3. Treating People as Means: The Supreme Emergency Exception in Rawls's JAF

Let me explore my reference to a tension and hence a possible unavoidable incoherence between Rawls's Kantian conception of JAF and his consequentialist conception

⁹ He embraces Michael Doyle's view that liberal states do not engage in war among themselves, but they often do so against nonliberal belligerent states (see Doyle 1997, esp. 474–84).

of a supreme emergency in a just war. An important reason why Rawls proposes and defends his conception of JAF is precisely to avoid treating people only as means, which he assumes to be a trademark of consequentialist reasoning, especially utilitarian reasoning. Instead, in the spirit of Kant, he argues for the need to respect people's moral integrity as ends rather than as means only because they have a rational capacity for moral autonomy. Rawls underlines that the distinctive feature of Kant's moral theory is his focus on rational autonomy rather than the generality and universality of his theory (Rawls 1999e, 221).

While Rawls insists that Kant's categorical imperative of humanity needs to be interpreted, he contends that treating people as ends "implies at the very least treating them in accordance with the principles to which they would consent in an original position of equality" (ibid, 156–7). For Rawls, such a hypothetical consent is necessary to act according to Kant's categorical imperative of humanity. So, if people were unable to act according to principles that they could consent to in a position of equality, such as his original position, they would be unable to act according to Kant's categorical imperative of humanity. However, according to Rawls, people in the original position behind a veil of ignorance are "free and rational persons concerned to further their own interests" (Rawls 1999e, 10, 123–4). That is to say, they are rational self-interested persons who are motivated by their own good rather than by other people's well-being or by a greater good. Hence, they are not altruistic. In addition, he stipulates that they are neither selfish nor envious.

Rawls operates with a contestable concept of rationality based on maximizing one's self-interest without necessarily being concerned with other people's interests. He assumes that his hypothetical choosers in the original position behind a veil of ignorance are rational self-interested individuals. Yet one might conceive of rational self-interest in at least two ways: in a narrow or in a broad sense. Rational self-interested people in the narrow sense are those who choose effective means to maximize satisfying their desires to further their own good ignoring that they might, at times, have a weightier obligation to promote a greater good. By contrast, rational self-interested people in the broad sense are those who choose effective means to maximize satisfying their desires to further their own good only if, in doing so, they do not neglect their weightier obligation to promote a greater good (Lemos 1986, 30–44).

Behind the veil of ignorance, Rawls's hypothetical choosers operate with the narrow rather than with the broad conception of rational self-interest to guarantee that they will opt for his two principles of justice. That is so because they do not possess any knowledge that might skew their choices. In addition, they are ignorant of what position they would occupy in actual political society once the veil is lifted. They want to make sure that, despite his principles of justice being impartial, each chooser, being a rationally self-interested person in the narrow sense, seeks her own good without necessarily being concerned with the good of others or with a greater good.

As result, Rawls faces a dilemma. Either he holds onto his narrow conception of rational self-interest that operates behind the veil of ignorance or, once the veil is lifted, he could adopt the broad conception of rational self-interest or one akin to it. If so, Rawls's rationally self-interested persons in the narrow sense, be they behind a veil of ignorance or in actual political society, could not consistently embrace his self-interested narrow

conception of rationality and Kant's categorical imperative of humanity, and simultaneously hold the following principle: It is right to target objectively innocent but rationally self-interested people in the narrow Rawlsian sense who might be living under an authoritarian aggressive state if that is for a greater good and is the only way that such a vicious regime could be defeated during a supreme emergency exemption. However, once the veil is lifted, people in actual political society who adopt the broad concept of rational self-interest, but not necessarily Kant's categorical imperative of humanity, might consistently hold the above-mentioned principle.

As I see it, Rawls faces two options to try to avoid the already mentioned incoherence, none of which seems to have a felicitous outcome. Once the veil is lifted, he could have adopted the broad conception of rational self-interest, which might be compatible with his call to a supreme emergency. But even if he had done so, he would still have faced the problem of explaining away how people who embrace Kant's categorical imperative of humanity would choose and therefore consent to a principle of action whereby they, as well as others, will be treated as means only for a greater good during a supreme emergency.

A supreme emergency exception creates a tension for those who, like Rawls, embrace the narrow and/or the broad self-interested conception of rationality and Kant's categorical imperative of humanity—which values people as embodiments of rationality being ends in themselves—and simultaneously argue that during such an exception one can justify a policy of deliberately targeting objectively innocent people whose rationality would be destroyed, presumably for a greater good, such as winning a just war. That is, an appeal to winning a just war is insufficient to palliate the tension that exists between his two positions, namely, the narrow or the broad self-interested conception of rationality and upholding Kant's categorical imperative of humanity while sacrificing one's self-interest, humanity, and hence rationality during a supreme emergency exception to bring about a greater good.

Also, while Rawls's JAF presupposes that treating people as ends minimally implies treating them according to principles that they would hypothetically consent to in an original position of equality, the categorical nature of Kant's imperative of humanity does not depend on hypothetical consent but rather on the actual recognition of people's humanity as rational agents whose rational capacity is valuable in itself and thereby confers value on their choices (Korsgaard 1996, 122–3). It seems that to interpret Kant's categorical imperative of humanity as Rawls has proposed is to interpret it as a hypothetical rather than as a categorical imperative, which is inconsistent with Kant's view of morality.

For Kant, we treat people solely as a means for our own ends only when we intend to act towards them based on principles that they cannot possibly agree with, and when they cannot share the end of our action (Kant 1964, 97 [AK 4:430]). As he states when discussing the categorical impermissibility of a false promise, "he [the promisor] is intending to make use of another man *merely as a means* to an end he [the promisee] does not share." He continues: "For the man whom I seek to use for my own purposes by such a promise cannot possibly agree with my way of behaving to him, and so cannot himself share the end of the action" (ibid.). Thus for Kant lying is always morally wrong, which seems to be a rather rigorist position. However, his objection is not only based on whether in the act of lying we treat people solely as a means for our own ends, because we could be lying for their own good. He also objects to the act of lying because those who we lie to are prevented from exercising their rational autonomy or moral agency.

Moreover, the person whom I intend to lie to, even if I were to do so for her own good, could never agree with my action because it would be actually impossible for her to know the illocutionary content of my act of lying. Similarly, she would be prevented for sharing the end of my action. We can always conjecture that if the person we are lying to were to be in our position, she would have acted likewise. But that person is not in our position. Hence, we cannot validly conclude that she would have acted likewise. Whether she could have chosen to act likewise remains an open question.

Given the above open question, it stands to reason that for Kant lying is always morally impermissible because of our disrespect of those whom we lie to as well as for the same disrespect for our own rational autonomy. According to him, “the greatest violation of man’s duty to himself regarded merely as a moral being [...] is the contrary of truthfulness, *lying*” (Kant 1991a, 225 [AK 6:429]). Also, he offers an argument defending one’s categorical duty to preserve oneself. As a result, he argues against deliberately committing suicide and/or self-mutilation, precisely because “to annihilate the subject of morality in one’s own person is to root out the existence of morality from the world, [...] even though morality is an end in itself” (ibid., 219 [AK 6:423]). Moreover, since he conceives of suicide as “willfully killing oneself,” he calls such an act “murdering oneself (*homocidium dolosum*)” (ibid., 218 [AK 6:422]). Given, therefore, that murder is categorically wrong, it necessarily follows that for him suicide is also categorically wrong.¹⁰

Some scholars argue against such a rigorist interpretation of Kant’s categorical imperative of humanity. For example, Allan Wood argues that every argument based on the categorical imperative of humanity needs a logically independent second premise to validly conclude how one ought to act in a given situation. Assuming that we ought to act according to the categorical imperative of humanity, one may conjoin it with a second premise the content of which is as follows: The deliberate killing of oneself violates one’s humanity as an end in itself. Hence, one ought never to deliberately kill oneself.

According to Wood, since the second premise is logically independent from the categorical imperative of humanity, and the conclusion validly follows only if we assume the truth of the second premise, while at the same time assuming the truth of the categorical imperative, one might reject the truth of the conclusion by rejecting the truth of the second premise. Even if one were to agree with Wood’s nonrigorist interpretation of Kant’s categorical imperative, it would be most difficult to conceive of how one might accept the categorical imperative of humanity and simultaneously argue that there are instances under which the deliberate killing of oneself does not violate one’s humanity as an end in itself (Wood 1998, 180–5).

Similarly, assuming that we ought to act according to the categorical imperative of humanity, one may conjoin it with a second premise the content of which is as follows: The deliberate targeting of innocent noncombatants always fails to respect as end in themselves the humanity of those who have been targeted. Hence, we ought never to engage in the deliberate targeting of innocent noncombatants. That is not to deny that under exceptional circumstances, such as a supreme

¹⁰ For a compelling defense of Kant’s rigorist opposition to suicide or self-disembodiment under any circumstances, see Unna 2003.

emergency situation, one might argue that a greater good, as Rawls allows in his just war paradigm, trumps Kant's categorical imperative. If that were to be the case, then one's position would not be Kantian, as Kant understood his position, but rather consequentialist.

Despite Wood's challenging interpretation of the categorical imperative of humanity, one might argue that his interpretation is not true to the letter or spirit of Kant's views. For example, Kant argues that we have a perfect negative duty against the act of suicide for two mutually supportive reasons. First, because, for him, a good will is intrinsically good (or good without qualification); and second, in the act of suicide we are deliberately destroying the rational autonomy that makes morality possible, namely, we are deliberately destroying a good will. Therefore, he seems to assume, as Wood suggests, that we do not need a second premise to understand that suicide is categorically wrong. For Kant, it is evident that one who deliberately commits suicide is failing to respect her humanity as an end and is treating her humanity solely as a means to end her life. Whether one agrees with Kant or not, he seems to embed his concept of morality into his concept of humanity, and vice versa.

Kant argues that whoever violates his categorical imperative is incoherent in the sense that the person's behavior might be described as paradoxical. In his argument against suicide, he tries to establish such a paradox. For example, consider the following statement: "I am a rationally autonomous person," (namely, my rational autonomy or good will is what makes morality possible and hence it is intrinsically good). Next, consider the following statement: "I deliberately choose to commit suicide" (namely, I choose to destroy that which makes morality possible in the first place). My position is paradoxical, namely, unreasonable, because my two beliefs are mutually exclusive in Kant's sense. A similar paradox would arise when one upholds the categorical imperative of humanity and deliberately chooses to target objectively innocent noncombatants in a just war. That is to say, one values the rational autonomy of people which makes morality possible, but simultaneously contends, as Rawls argues, that it is right, in the sense of being justified, to deliberately target objectively innocent people in a just war during an emergency situation. From the perspective of Kant's categorical imperative of humanity, Rawls's position is paradoxical.

I grant that Kant's view of morality is rigorist and that the above conclusions follow validly from his categorical imperative because the single-premise argument is circular. But all deductively valid arguments, whether they have one or more premises, are always to some extent circular. The relevant question is whether the validity is vicious or illuminating. I think that, in the case of his categorical prohibition of suicide, and similarly with my example of the categorical prohibition of deliberately targeting objectively innocent noncombatants, the validity is illuminating.

Kant categorically objects to violating one's perfect duties to oneself because under no condition would he justify lying or deliberately killing oneself. In both instances we would be destroying our rational moral autonomy. If that is so, then it would necessarily follow that, in a supreme emergency in a just war, he would categorically object to the deliberate targeting of objectively innocent noncombatants whose rational autonomy is consequently destroyed, and would object whether the targeting is for a greater good or for a lesser evil. By contrast, Rawls writes, "since

(let's suppose) there are no absolute rights [...] there are occasions when civilians can be attacked directly by aerial bombing" (Rawls 1999b, 568). As result, he justifies the deliberate killing of civilians, who are oftentimes objectively innocent, and hence whose rational autonomy is deliberately destroyed if a greater good or a lesser evil is accomplished.

To illustrate the above point, consider how Rawls defends Great Britain's policy of area bombing against the civilian population of Nazi Germany at the beginning of WWII, and only then because the British faced a supreme emergency situation. According to Rawls, "under no conditions could Germany be allowed to win the war" (ibid.; cf. Rawls 1999d, 99). He rightly suggests that the rise of Nazism created an extreme crisis for European nations. During such an anomalous situation, he claims, it is justified to area-bomb objectively innocent civilians, and he thereby simultaneously justifies violating the PNCI.

While in his 1995 article, under Principles 4 and 6, Rawls mentions the justified violation of the PNCI when faced with an extreme crisis, in his LOP he cites such justified violation in Principle 6, where he instead adopts the language of supreme emergency.¹¹ In both works, he offers the following two reasons to justify violating the PNCI: (1) the peculiar evil nature of Nazism, which threatened civilized society everywhere, and (2) the vulnerability of constitutional democracy as represented in European culture (Rawls 1999b, 569; 1999d, 99).

Rawls's appeal to the commendable nature of constitutional democracies broadly construed to justify violating the PNCI needs to be qualified. While constitutional democracies have contributed to the development of positive institutions and continue to do so in different parts of the globe, as by promoting democratic institutions and the protection of fundamental human rights and the rule of law, its pedigree at times has been morally tainted. Rawls tacitly recognizes this problem when he acknowledges that a well-ordered people can have its hands dirty and "still have the right and even the duty to go to war to defend itself. This is clear from the history of World War II" (Rawls 1999d, 94 n. 7).

The issue, however, is not whether, for example, the British had the right or duty to go to war against the Nazis, who threatened their sovereignty and hence their parliamentary democracy. Of course, they did have the right and duty to protect their citizens and institutions against Nazi aggression. Nevertheless, the controversial moral issue is whether the British, or anybody else facing such a threat, could be morally justified in adopting a policy of area bombing against the aggressor's civilian population.

At a practical level, even during a supreme emergency exemption, the practice of area bombing, which is just a euphemism for the practice of emergency terrorism, seems to violate the PNCI embedded in contemporary *jus ad bellum* and *jus in bello* restrictions in IHL. If that is so, one could argue that such a practice calls into question the nature of a just war. Those who engage in emergency terrorism against objectively innocent noncombatants are doing something gravely wrong by deliberately harming innocent

¹¹ Since in LOP Rawls admits that his view on just war was influenced by Walzer's 1992 *Just and Unjust Wars*, where Walzer uses Winston Churchill's expression "supreme emergency," that might explain why Rawls changed his usage, replacing the expression "extreme crisis" in Rawls 1995 with the expression "supreme emergency" in LOP. For Walzer's conception of supreme emergency, see Walzer 1992, 251–8.

civilians. Such a harm appears to violate Rawls's fifth principle or condition for a just war. He, like Kant, recognizes that the way we fight a war may or may not set the stage for future wars. Therefore, the practice of emergency terrorism might undermine an important motive for engaging in a just war, which is to secure a lasting peace.

By justifying the deliberate targeting of objectively innocent people during a supreme emergency exemption, Rawls appears to adopt a consequentialist approach. He assumes that, in such an emergency, those who are defending the values of a liberal constitutional democracy against the presumably existential threat posed by an authoritarian aggressive state are justified in targeting objectively innocent non-combatant not only as a matter of right but also as matter of duty for a greater good.

Regarding the concept of "extreme necessity," which seems to be equivalent to that of "supreme emergency," Rawls argues that to determine whether such an anomalous situation exists is not a matter of balancing interests but rather a matter of judgment "as to whether certain objective circumstances are present which constitute the extreme crisis exemption" (Rawls 1999b, 568 n. 7). While he acknowledges the vagueness of the concept in question, he underscores that its applicability "rests on judgment" (ibid.).

One can agree that whether an extreme crisis or a supreme emergency exists is a matter of objective judgment. For example, one may ask the following questions: Was Great Britain in 1940–41 facing a supreme emergency, namely, a real existential threat by Nazi Germany at the beginning of WWII? Did Imperial Japan in August 1945 represent a real existential threat to the Allies, especially to the United States? Rawls and others answer yes to the first question and no to the second.

Let us suppose that we answer yes to both questions and that, in addition, our judgments are compelling. The next question we need to address is: What policies could we justifiably implement in the face of such a supreme emergency or real existential threat? The answer to this question is not "only" a matter of judgment, as Rawls has suggested, but is also a matter of balancing interests. Ultimately, we are balancing the interests of those who defend the values embedded in liberal constitutional democratic states, namely, well-ordered peoples, above those who defend the values embodied in authoritarian aggressive states. Whether one wants to conceive of the "balancing" in terms of the protection of interests or of the protection of rights is simply incidental to whether we can come up with a reasonable and compelling judgment, which might be what Rawls had in mind. One can make the case that during a supreme emergency, numbers do matter in attempting to justify, permit, or excuse the targeting of objectively innocent noncombatants.

In the event of a supreme emergency, Rawls's argument allows for appealing to numbers or to the greater good sacrificing the categorical immunity of the objectively innocent not to be deliberately targeted. So, under this dire scenario, he justifies deliberately killing the innocent few to save the lives of the innocent many for a greater good. This practice is consistent with Rawls's conception of "telishment," namely, punishing the innocent few to save or improve the lives of the innocent many. Yet he evidently objected to the practice of telishment (Rawls 1999a, 27). One can plausibly argue that he developed his Kantian conception of JAF partly to avoid telishment. And yet, while it is true that he developed his conception of telishment as being applicable to his view of punishment rather than to a conception of just war, it is also true that in the just war tradition there is a sense in which those who are in the right are not only defending themselves but are also punishing the aggressors for their unjust behavior.

Rawls could not defend his Kantian conception of JAF, where he justifies treating each individual person as an end, and simultaneously operates with the narrow or the broad conception of rational self-interest, while appealing to consequentialist considerations whereby individual persons are treated as means only to bring about a greater good. If he were to do so, he would be incoherent. Moreover, incoherent beliefs are excluded from his different conceptions of the original position and from his conception of public reason. Otherwise, his fictional or real people might be unreasonable rather than reasonable, as Rawls assumes them to be.

4. Closing Remarks

In this paper, I have tried to demonstrate that my original concern about Rawls's fundamental tension between his ideal Kantian conception of JAF and his realist consequentialist conception of just war, rather than being a misguided concern or a purely rhetorical one, is a relevant philosophical and political concern about what appears to be an incoherence between his ideal and practical theories. Moreover, I have underscored that Rawls's justification for targeting objectively innocent people during a supreme emergency exemption could be described as emergency terrorism. But emergency terrorism treats people as means only. Rawls's JAF requires that they not be so treated. Therefore, there seems to be a substantive tension between his conception of JAF and his justification of emergency terrorism. This tension shows the limits of his ideal conception of JAF.

Those who are critical of my argument against Rawls could offer at least three plausible replies. First, they might underscore that Rawls never intended his theory of justice to deal with issues regarding terrorism. Nevertheless, his justification for deliberately targeting objectively innocent people during a supreme emergency exemption could be conceived of as emergency terrorism. As Rex Martin and David Reidy argue in discussing the similarities and differences between Walzer's and Rawls's conceptions of just war, "Perhaps the simplest point to make here is that traditional just war theory must be expanded to deal with issues of terrorism" (Martin and Reidy 2006, 14). By justifying what might be conceived of as emergency terrorism during a just war, Rawls justifies violating the PNCI. By doing so, he has already expanded or modified traditional and contemporary just war theory as expressed in LOAC or contemporary IHL.

Second, some might argue that there is no real tension between Rawls's Kantian conception of justice and his conception of a supreme emergency exemption. Since we have reason to believe that all doctrines, principles, or rules might have exceptions, especially during extraordinary circumstances, we could argue that taking people's rights seriously is not incompatible with deliberately killing objectively innocent noncombatants for the greater good during a supreme emergency. Still, an appeal to the greater good seems to be incompatible with Rawls's Kantian conception of JAF and his criticism of *telishment* where he argues against trading off people's rights for consequentialist considerations.

Third, some might argue that during a supreme emergency the stakes are so high that only then could the deliberate violation of objectively innocent people's rights be justified for the sake of promoting a greater good, such as protecting a well-ordered people. Rawls could justify the targeting of objectively innocent noncombatants during a supreme emergency because doing so is likely to bring about a greater good,

as would happen by defeating authoritarian aggressive states like Nazi Germany or Imperial Japan. Yet doing so is equivalent to treating objectively innocent victims as means only, rather than as ends. A greater good might excuse, but not necessarily justify, the targeting of innocent noncombatants, as Rawls has argued, especially in casting his argument in favor of JAF and against telishment.

Rawls's justification for deliberately targeting objectively innocent people during a supreme emergency seems more realistic, and hence non-Kantian and consequentialist, than utopian. On the contrary, his insistence on the reasonableness of JAF in TJ, PL, and LOP, especially his Kantian conception of respecting people as ends rather than as means only, seems more utopian than realistic. I am afraid, therefore, that his Kantian conception of JAF is irreconcilable with his conception of a supreme emergency exemption. As a result, in light of his ideal Kantian conception of JAF, his realistic and consequentialist conception of a supreme emergency within his just war paradigm seems to be incoherent and hence morally flawed.

Department of Philosophy
Seton Hall University
400 South Orange Ave
South Orange
NJ 07079
United States
Email: medinavi@shu.edu

References

- Aquinas, T. 2006. *Summa Theologiae*. Vol. 38, Injustice (2a2æ. 63–79). Cambridge: Cambridge University Press.
- Doyle, M. 1997. *Ways of War and Peace: Realism, Liberalism, and Socialism*. New York: W. W. Norton & Company.
- Kant, I. 1964. *Groundwork of the Metaphysics of Morals*. Trans. H. J. Paton. New York: Harper & Row.
- Kant, I. 1991a. *The Metaphysics of Morals*. Trans. M. Gregor. Cambridge: Cambridge University Press.
- Kant, I. 1991b. Perpetual Peace: A Philosophical Sketch. In *Political Writings*. Ed. H. Reiss, 93–130. New York: Cambridge University Press.
- Korsgaard, C. M. 1996. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- Lemos, R. M. 1986. *Rights, Goods, and Democracy*. Newark: University of Delaware Press.
- Martin, R., and D. A. Reidy, eds. 2006. *Rawls's Law of Peoples: A Realistic Utopia?* Malden, MA: Blackwell Publishing.
- Medina, V. 2010. Militant Intolerant People: A Challenge to John Rawls' Political Liberalism. *Political Studies* 58(3): 556–71. <https://doi.org/10.1111/j.1467-9248.2009.00797.x>.
- Medina, V. 2015 *Terrorism Unjustified: The Use and Misuse of Political Violence*. Lanham, MD: Rowman & Littlefield.

- Orend, B. 2000. *War and International Justice: A Kantian Perspective*. Waterloo, ON: Wilfrid Laurier University Press.
- Pogge, T. 2006. Do Rawls's Two Theories of Justice Fit Together? In *Rawls's Law of Peoples: A Realistic Utopia?* Ed. R. Martin and D. A. Reidy, 206–225. Malden, MA: Blackwell Publishing.
- Rawls, J. 1995. 50 Years after Hiroshima. *Dissent* (Summer): 323–7. <https://www.dissentmagazine.org/article/50-years-after-hiroshima-2/>.
- Rawls, J. 1996. *Political Liberalism*. New York: Columbia University Press.
- Rawls, J. 1999a. Two Concepts of Rules. In *Collected Papers*. Ed. S. Freeman, 20–46. Cambridge, MA: Harvard University Press.
- Rawls, J. 1999b. Fifty Years of Hiroshima. In *Collected Papers*. Ed. S. Freeman, 565–72. Cambridge, MA: Harvard University Press.
- Rawls, John. 1999c. The Idea of Public Reason Revisited. In *Collected Papers*. Ed. S. Freeman, 573–615. Cambridge, MA: Harvard University Press.
- Rawls, J. 1999d. *The Law of Peoples*. Cambridge, MA: Harvard University Press.
- Rawls, J. 1999e. *A Theory of Justice*. Rev. ed. Cambridge, MA: Harvard University Press.
- Rawls, J. 2001. *Justice as Fairness: A Restatement*. Ed. E. Kelly. Cambridge, MA: Harvard University Press.
- Roberts, P. 2011. The Supreme Emergency Exemption: Rawls and the Use of Force. *European Journal of Political Philosophy* 11(2): 155–71. <https://doi.org/10.1177/1474885111425120>.
- Unna, Y. 2003. Kant's Answers to the Casuistical Questions Concerning Self-Disembodiment." *Kant-Studien* 94(4): 454–73. <https://doi.org/10.1515/kant.2003.026>.
- Vitoria, F. de. 1991. On the Law of War. In *Political Writings*. Ed. A. Pagden and J. Lawrance, 293–327. Cambridge: Cambridge University Press.
- Walzer, M. 1992. *Just and Unjust Wars: A Moral Argument with Historical Illustrations*. 2nd ed. New York: Basic Books.
- Walzer, M. 2004. Emergency Ethics. In *Arguing about War*, 33–50. New Haven, CT: Yale University Press.
- Williams, H. 2012. *Kant and the End of War: A Critique of Just War Theory*. New York: Palgrave MacMillan.
- Wood, Allen. 1998. Humanity as an End in Itself. In *Kant's Groundwork of the Metaphysics of Morals: Critical Essays*. Ed. P. Guyer, 165–187. Lanham, MD: Rowman & Littlefield.