

*Culpable Bystanders, Innocent Threats and the Ethics of Self-Defense*¹

YITZHAK BENBAJI
Bar-Ilan University, Ramat-Gan 52900
and
Shalom Hartman Institute
Jerusalem 93113
Israel

The moral right to act in self-defense seems to be unproblematic: you are allowed to kill an aggressor if doing so is necessary for saving your own life. Indeed, it seems that from the moral standpoint, acting in self-defense is doing the right thing. Thanks, however, to works by George Fletcher² and Judith Thomson,³ it is now well known how unstable the

1 The first version of this paper was presented in the 2004 meeting of the Middle East Legal Studies Seminar, sponsored by Yale Law School. I would like to thank Anthony Kronman, Owen Fiss, Kenneth Mann, and Avishai Margalit for the opportunity to present the paper and for their helpful comments. The comments by other participants in the seminar were very helpful as well. Let me mention just few: Sharhabeel Al Zaeem, Guido Calabresi, Hanoach Dagan, Nissreen Haram, Reva Siegel, and Robert Post. Later versions of this paper were presented at the conference 'War, Terrorism and Rogue States' sponsored by Bar-Ilan Law School and in the research group of Bar-Ilan philosophy department. Comments by Yuval Dolev, Shachar Lifshits, Charlotte Katzoff, David Widerker, and Yafa Zilbershatz were very valuable. Finally, I am particularly indebted to Ron Shapira, Jeff McMahan, Ira Schnall, Daniel Statman, Michael Walzer, Noam Zohar, and the Editors of the *Canadian Journal of Philosophy* for detailed written comments in light of which the present version of this paper was shaped.

2 G. Fletcher, 'Proportionality and the Psychotic Aggressor,' *Israel Law Review* 8 (1973) 367-90

3 J. Thomson, 'Self-Defense,' *Philosophy and Public Affairs* 20 (1991) 283-310; cf. also a much earlier paper, in which Thomson extensively uses the notion of Self-Defense: 'A Defense of Abortion,' *Philosophy and Public Affairs* 20 (1971) 41-66.

moral basis of the right to self-defense is. We are in the dark with regard to one of the most basic problems raised by this right, namely: the problem of the innocent aggressor. The disturbing question is simple enough: Is a potential victim allowed to kill, in self-defense, a morally innocent aggressor?

Thomson suggests an answer to this question, and provides a powerful argument in support of it. According to her, if the aggressor holds no right to kill you (i.e., his aggression is not a response to a lethal threat you pose to him), you may kill him *just because* otherwise he would kill you.⁴ Hence, killing an aggressor is morally justified, even if he is free of fault in committing the particular aggression that threatens your life. Culpability has nothing to do with the right to self-defense, since an agent who exercises this right is not an agent of justice. The major argument Thomson advances appeals to the presumption against exercising private violence (I call this presumption the Weberian Presumption). She takes it as almost self-evident that the Presumption cannot be defeated either by the aggressor's being *liable* to harm, or by his *deserving* it. (I shall explain the liability / desert distinction later on.) She infers that it is defeated simply because the aggressor is about to kill someone and has no right to do so. In a nutshell, this is the conclusion of what I call Thomson's 'Weberian Argument.'

It is clear, however, that Thomson fails to prove the strong claim she infers from the Weberian Presumption, namely, that culpability is wholly irrelevant to the right to self-defense. The ethics of self-defense against innocent threats and the ethics of self-defense against villainous threats are asymmetrical.

I aim in this paper to explain this asymmetry in a new way. I shall first claim that a new explanation is needed: the Weberian Argument does cast serious doubt on the justice-based explanation of the asymmetry advanced by Jeff McMahan and others (following McMahan, I call his explanation 'Justice.') Second, I shall argue for a right-based ethics of self-defense. The asymmetry can be explained within this framework due to a special interest that a victim has in stopping a villain from killing him, an interest that he does not have if the aggressor is innocent. Specifically, when one defends oneself against an innocent threat, one avoids (more or less) harmful death. In defending oneself against a villainous threat, one avoids *disrespectful* death. This is why culpability matters.

I shall begin with a critical presentation of Thomson's solution to the problem of the innocent threat and distinguish two arguments that she

4 See Thomson, 'Self-Defense,' 289.

gives for it, the Intuitive Argument and the Weberian Argument (section I). I shall then show that the Intuitive Argument should be modified in order to capture our intuitive responses to innocent-threat cases. It should be modified in a way that strongly suggests an interest theory of the right to self-defense against innocent threats (section II). Next, I show that the Weberian Argument fails: there is a sharp intuitive asymmetry between innocent threats and culpable aggressors regarding the scope of the right to kill them in self-defense (section III). Then, I shall reject the explanation provided by justice-based accounts for this asymmetry, by showing that, appropriately reinterpreted, Thomson's Weberian Argument refutes those accounts (section IV). In section V, I defend a right-based — in fact, an interest-based — explanation of why culpability matters in the ethics of self-defense. Finally, I argue that, according to this explanation, even if one possesses a right to kill an innocent aggressor in self-defense, this might well be (what Waldron calls) a moral right to do wrong, i.e., a right to do what is all-things-considered morally *unjustified* (section VI).

I The Right to Self-Defense: Thomson's Theory Presented

Thomson's basic claim is simple: you have the right to kill an aggressor in self-defense if he would otherwise kill you (while he holds no right to do so). Indeed, you have this right, *because* otherwise the aggressor would kill you. But what if the person whose being killed is necessary for preventing your death became an aggressor through no fault of his own? For example, what if the aggressor hallucinates that you are about to kill him? What if he is under the influence of drugs that cause him to be extremely aggressive? Furthermore, imagine that the one who is about to kill you is not an aggressor at all. You are sitting on a bench, and due to some cause, a man is falling on you — someone pushed him, or he lost his balance, or whatever. It turns out that, unfortunately for both of you, unless you blow him to pieces with a gun you have handy, he will crush you. Do you have the right to defend yourself against such an *innocent threat*? Thomson answers affirmatively. You are allowed to kill anyone who would otherwise kill you. True, the falling man won't kill you by any *action* of his, but still, he would kill you the way a stone can kill you — by falling on you.⁵ And he has no right to do so.

5 Ibid., 287-9

As Thomson points out, one important advantage of her theory is that it does not allow you to kill (what is called) a bystander in self-preservation. In her view, self-defense turns out to be a drastically circumscribed right. Indeed, there is an obvious *factual* difference between the 'Yes cases' described above (the cases of villainous or innocent aggressors and the case of the innocent threat), cases in which it is justified to kill in self-defense, and a 'No case,' in which you shield yourself from the bullets that are about to hit you by using a bystander who just happened to be around. In the latter case, you defend yourself by killing a person who does not threaten you, because it is the only way for you to save your life. Thomson's theory implies that you are not allowed to kill a bystander, not because the bystander is innocent — after all, the innocent threat is also innocent — but rather because he is merely a bystander.⁶ You may kill an aggressor or a threat, even if he is innocent, in order to prevent *him* from *killing* you; but you may not kill an innocent bystander simply in order to avoid your own death or to avoid your being killed by someone else.

Thomson insists, in addition, that her theory implies that 'the permissibility of X's killing Y in self-defense goes hand in hand with the permissibility of Z's killing Y in defense of X.'⁷ Indeed, this theorem, that the right to kill in self-defense implies a permission to kill in other-defense, follows from conceiving of self-defense as *justification* rather than as *excuse*. If X is justified in killing Y, there is a fact *about* Y that justifies X in killing Y. And this fact would be a justification for anyone to kill Y: X's justification to act in self-defense is essentially sharable. In Paul Robinson's words, 'since justification is dependant not on the actor but only on the act, an act which is justifiable when performed by one actor is necessarily justifiable when performed by any actor.'⁸ Excuses, on the

6 Ibid., 289 ff.

7 Ibid., 306

8 P. Robinson, 'A Theory of Justification,' *UCLA Law Review* 23 (1975) 266-92, at 278, n. 49; quoted in R. Christopher, 'Self-Defense and Defense of Others,' *Philosophy & Public Affairs* 27 (1998), 138. Note, however, that the claim attributed here to Thomson is not as general as Robinson's claim. Thomson's theorem regards justified killing in self-defense, while Robinson makes a general statement about justification. As a referee for the *Canadian Journal of Philosophy* correctly pointed out to me, many justifications are essentially agent relative. Suppose X is an official executioner for the government, and Y has been duly sentenced to death. X is *justified* in killing Y, but no one else is. Or, to take another example, you are justified in punishing your son, but no one else is. Two points should be stressed here: as far as I can see there cannot be essentially agent-relative justification for killing in self-defense. Secondly, even an agent-relative justification is sharable in some sense. With respect to the second example: I am *prima facie* justified in helping you to punish your son.

other hand, are essentially agent-relative. If X has an excuse for killing Y, it is because X is frightened, or coerced, or the like. In the paradigmatic cases of excused actions, X acts wrongly, but is not to be accused for his wrongdoing, since he was not in full control — *his* life was in danger (say), and he instinctively acted to save it. Thus, X is excused because of a certain mental state that no one else necessarily shares; hence, it does not follow from the fact that X is excused in harming Y, that anyone else is excused in doing so.

What are the arguments that support Thomson's doctrine of self-defense? How can it be that a person loses his right to life through no fault of his own? Some of Thomson's critics believe that her claims about self-defense are groundless, nothing more than 'an ingenious exercise in begging the question.'⁹ Contrary to these philosophers, I believe that Thomson does have a powerful argument in support of her conviction that it is permissible to kill an innocent threat in self-defense, an argument which is free of vicious circularity.

The first argument she advances — let us call it the Intuitive Argument — is short and effective. Suppose the innocent aggressor (or threat) is about to run you over with a truck. If killing an innocent aggressor in self-defense is wrong, acting in self-defense is merely *excusable*. And to say this is

... to say that although you would not be at fault for blowing up the truck in that case, you ought not to blow it up, you act wrongly if you do. I think that cannot be right. (I think it an excessively high-minded conception of the requirements of morality.)¹⁰

One way to put Thomson's thought here is the following. The intuition that killing the innocent aggressor in self-defense is permissible (that is, not wrong) is too central to be given up. Its centrality promises that it would be present in any set of principles that satisfies the Rawlsian requirement of reflective equilibrium. In light of the no less central intuition that killing a bystander in self-preservation is prohibited, the

9 J. McMahan, 'Self-Defense and the Problem of the Innocent Attacker,' *Ethics* 104 (1994) 252-90, at 278. McMahan's reason for this strong claim is that Thomson rejects the classical notion of rights exemplified by Raz's definition (Joseph Raz, *The Morality of Freedom* [Oxford: Clarendon Press 1986], 186). In Raz's view, one holds a right if an interest of his constitutes a sufficient reason for holding some other person under a duty.

10 Thomson, 'Self-Defense,' 285. One element in this argument will be rejected in the final section, namely: I shall agree that victim is not merely excused, but this does *not* imply that the victim is justified.

best interpretation of the permission to kill an innocent threat would appeal to the fact that if you don't kill the innocent aggressor, he will kill you, though he has no right to do so. Put in terms of duties: the innocent aggressor forfeits his right not to be killed because he is about to violate his duty not to kill you.

Let me now examine Thomson's Weberian argument. In its simple, unqualified form, this argument has three premises: first, intuitively, killing a villainous aggressor in self-defense is justified. Second, private persons are not agents of punitive violence; they should not exercise punitive coercive power. Third, had the right to kill in self-defense been dependent on the fault of the aggressor, it would be a form of punishment. It follows that, rather than being at fault, some other characteristic of the villainous aggressor is the basis of the right to kill him in self-defense. The most likely alternative candidate is simply the fact that if the victim does not kill the aggressor, the villainous aggressor would kill him. This is the most plausible candidate, in view of the equally strong moral intuition that killing bystanders in self-preservation is unjustified; for the bystander would not *kill* the victim. Thus, the dividing line between the Yes cases and the No cases is clearly delineated.

This argument is put forward when Thomson disapprovingly raises the possibility that the right to kill in self-defense is limited to killing villainous aggressors, based on the fact that the villainous aggressor (who in her story is a driver) deserves punishment for his aggression, whereas the fault-free driver does not. Here is what she has to say against this proposal:

But who are you, private person that you are, to be dishing out punishment to the villainous for the things that they do? And anyway, what makes it permissible for you to blow up the truck is not the fact that the driver in that case deserves punishment ... or else it would be permissible for you to blow up the truck even if you do not need to do so to save your life.¹¹

Circumstances in which self-defense is justified are such that the Weberian Presumption that only a state may legitimately exercise violence is defeated. Thomson believes that, had the aggressor's culpability been the defeater of the Presumption, culpability would defeat the presumption against unauthorized punishment not only in cases of self- or other-defense, but in most other cases as well.

Notably, the conclusion of the Intuitive Argument is not as far-reaching as the conclusion of the Weberian Argument. The former concludes

¹¹ *Ibid.*, 289

that you are allowed to kill a threat or aggressor because otherwise he would violate your right not to be killed by him. The latter concludes that the *only* reason you are allowed to kill the *villainous* aggressor is that otherwise he would kill you. That is, the Weberian Argument purports to show that the fact that he is a villain has nothing to do with the right to kill him in self-defense. Culpability, Thomson seems to argue, is irrelevant.

Before rejecting this strong thesis, it should be noticed that Thomson's arguments make explicit a well-entrenched conception of statehood, law, and coercive power: when private people are allowed to exercise violence, they are allowed to do so in order to defend their legitimate self-interests. The state allows it, because it cannot deny people the 'natural' right to defend what is most important to them against those who intentionally or unintentionally threaten their basic interests. This permission is not intended to bring about justice: the weight of the right to defend an interest is determined by how basic the interest is, rather than by considerations of justice. When one's life is at stake, one's most basic interest is at stake, and hence, subject to constraints of necessity and proportionality, one is allowed to do anything against those who threaten it. This conception of self-defense is an element of a larger libertarian picture, according to which the right of self-ownership is nearly absolute.

II Amending the Intuitive Argument

The conclusion of the Intuitive Argument should be importantly modified; as it stands it is simply false. Or so I shall argue in this section. Before doing so, I shall mention three objections, which, to my mind, do not really threaten the Argument.

Some philosophers find the resulting theory of rights intuitively weird — the Intuitive Argument, in particular, leaves them cold. How can it be — it is often asked — that a person violates his duty without doing anything? The falling man would kill you, in the sense that a piano might kill you, but it sounds silly to say of a piano that it infringes the right you have against it that it not threaten your life.¹² Unlike these philosophers, I do not find any genuine difficulty here. For example, I violate my duty to stay out of your apartment even if I was brought there while I was

12 N. Zohar, 'Collective War and Individualistic Ethics: Against the Conscripton of "Self-Defense,"' *Political Theory* 21 (1993) 606-22, at 608-9. Cf. M. Otsuka, 'Killing the Innocent in Self-Defense,' *Philosophy and Public Affairs* 23 (1994) 74-94.

sleeping. This is because, as opposed to pianos, I am subject to certain duties by virtue of my being an agent. It does not follow that violation of my duties involves agency. After all, I can, unknowingly, unintentionally, and even without doing anything, violate my duty to pay taxes, for example. It might be further argued, however, that agency is *analytically related* to any verb like 'violate'; that is, when a person is *now* 'in violation of his duty,' this means minimally that by virtue of that duty, there is something this person is *now* required to, and therefore, necessarily, can do. But I fail to see the force this further argument. Usually, when a person is *fully excused* in violating a duty, it is because he cannot avoid violating the duty; this is true, in particular, in the case of a psychotic aggressor. There are, of course, differences between a fully excused intentional attacker and the falling man, but these differences seem irrelevant to the question of whether one can be in violation of a duty even if one cannot do otherwise.

Another common worry is that the causal involvement that distinguishes threats from bystanders is terribly vague. In many cases (listed by Zohar¹³) the question whether Y threatens X is completely undecidable. Imagine someone who passively impedes your escape from a threat. Is he part of the threat or not? Zohar is right, I think, in claiming that there is no decisive answer to this question. But this does not cast doubt on Thomson's Intuitive Argument. The Argument shows that the distinction between aggressors and threats on the one hand, and bystanders on the other is morally relevant, and depends on the difference in the causal relations of them to the inevitable harm the victim might suffer. Like many other morally important distinctions (the person/non-person distinction is an obvious example), the threats/bystanders distinction is vague, so that in the borderline cases our *moral* intuitions are not clear. In fact, the unclarity of our moral intuitions supports Thomson's distinction. It shows that these moral intuitions depend on our causal intuitions — the former intuitions are unclear *because* the latter intuitions are unclear.¹⁴

The most typical complaint against the conclusion of the Intuitive Argument is that it fails to motivate the moral distinction between innocent threats and innocent bystanders. All the Intuitive Argument does is point out a *factual* difference, which has no *moral* significance: the

13 Zohar, 'Collective War and Individualistic Ethics,' 611-13

14 Zohar further claims that our moral judgment determines the causal relation rather than flows from it. See N. Zohar, 'Innocent and Complex Threats: Upholding the War Ethic and the Condemnation of Terrorism,' *Ethics* 114 (2004) 734-51.

victim has a right to kill the innocent threat, in the interest of self-preservation (because otherwise the innocent threat would kill him). But by the same token, unless the victim kills the bystander blocking his escape, the victim will be killed; why, then, doesn't this fact justify killing the bystander in the interest of self-preservation? The Argument does not provide any explanation as to why the falling man's causal relation to the harm (which the victim is about to suffer) makes such a dramatic moral difference. This is a fair complaint, and I shall deal with it in greater detail later. For now, suffice it to say that most theories of self-defense take causal relations to be relevant, indeed, crucial to the justification of acting in self-preservation. Hence, they all face the above worry. Thomson's notion of right, though far from being free from difficulties, provides a plausible solution: the moral difference between the innocent threat and the bystander is that the threat, but not the bystander, violated his duty not to threaten the victim's life.

It seems that, generally, Thomson's opponents believe that we have to resist the Intuitive Argument's conclusion, *despite the fact* that most people believe it. Even to her opponents, her view with regard to innocent threats seems to be compelling. I shall argue, against this piece of common wisdom, that the conclusion of the Intuitive Argument should be importantly modified in order to capture our commonsense intuitions. Intuitively, in some circumstances, the non-threatening victim has no right to kill an innocent threat in self-defense, even if otherwise the innocent threat would kill the victim and thereby violate a right he holds against him.

To show this, let me point to a particularly revealing unacceptable consequence of the *Weberian* Argument. Surprisingly, it implies that, in some contexts, the numbers should not count. To see why, consider a case where eight villain aggressors are attacking you. It seems self-evident that if the only way to protect yourself is by killing all of them, you are allowed to do so. But, according to Thomson, the fact that they are villainous must be irrelevant. You are justified in killing them, *just* because otherwise they would kill you while holding no right to do so. To quote again, 'who are you, private person that you are, to be dishing out punishment to the villainous for the things that they do?'

But this has unacceptable consequences. If the factual basis for your right to kill in self-defense is *just* that the eight aggressors would otherwise kill you (while holding no right to do so), then you should equally have the right to kill eight *innocent* aggressors; thus you are justified in blowing up an elevator which is falling on you, however large the number of innocent people trapped inside it. Moreover, even a third party is justified in blowing up the elevator and all the people inside it in order to save you. This strongly suggests that Thomson's formula deduced from the Intuitive Argument (that X has a right to kill Y in

self-defense, if otherwise Y would kill him and thereby violate the right X holds against Y) should be qualified: X is allowed to kill in self-defense only if there is only one (two? three?) innocent aggressor(s) who would otherwise kill him. Our intuitions do not tell us exactly how many innocent threats we may kill in self-defense; it seems clear, however, that if the number is large enough, one loses the right to defend oneself from innocent threats.

Even as thus qualified, Thomson's conclusion from the Intuitive Argument is wrong. To see why, let us consider a slightly varied case of innocent threat. The victim is sitting on a bench, when he sees a man falling on him. As before, the victim has a gun by which he can blow the falling man to pieces. If the victim does not do so, the falling man will kill him. But now there is a crucial difference: it so happens that by using his gun the victim would activate a bomb, located just next to the victim, whose explosion would certainly kill him (right after he kills the falling man.) Most people would agree that in such a case, when by killing the innocent threat the victim gains nothing but another two or three seconds of life, the victim is not allowed to kill him. This is so, despite the fact that if the victim does not kill the falling man, the falling man would kill him and thereby violate a right the victim holds against him. If this common opinion is right, Thomson's formulation has to be amended — it is not the case that what gives the victim the right to kill the falling man is simply that otherwise the falling man would kill the victim.

The relationship between these two examples should be obvious. Together they show that, intuitively speaking, before attributing to X the right to kill innocent people in self-defense, we must determine, on the one hand, how 'costly' X's action is in terms of lives of innocent people, and, on the other, how beneficial to X this killing would be. The right to kill an innocent threat in self-defense is sensitive to factors whose relevance Thomson's formulation either does not mention or outright denies.

The Intuitive Argument does not commit Thomson to rejecting these intuitive truths. (As we shall see, the Weberian argument, unfortunately, does.) She can embrace them by invoking the proportionality constraint: the fact that someone is about to violate a right that you have is not sufficient to justify your *killing* that person. In general, the fact that a person will otherwise violate your right not to be killed makes this person liable to defensive action, but this defensive action must also be proportionate. In the elevator case (where the number is large enough) and in the case in which the defender would die anyway very shortly after engaging in self-defense, killing in self-defense is not a proportional response.

In fact, it is pretty obvious that Thomson would welcome the second amendment. To show this, I shall develop an argument that proves that

X's right to self-defense is sensitive to the extent to which his premature death would be harmful to him.

Thomson discusses cases in which Y is innocently threatening X, but *not* vice versa. Thanks to this asymmetry, she can claim that X has the right to kill Y in self-defense, but Y has no right to fight back. But if what makes killing the falling man justified is that otherwise he would kill someone, it is easy to imagine symmetrical cases.¹⁵ For instance, suppose X and Y are driving two trains in opposite directions on the same track — they started driving at the same time, and a total-loss crash in which both sides are killed is bound to happen, unless one side stops the other by using deadly force. Now, clearly, a third party is not allowed to kill both X and Y. It would be silly for Z to argue that he is allowed to kill X in order to stop X from killing Y *and* to kill Y in order to stop Y from killing X. Yet, without the second correction suggested above, Thomson's theory, conjoined to other plausible assumptions, would imply that this is exactly the case.

I shall construe an argument that should be conceived as a *reductio ad absurdum* of the formulation that X has a right to kill Y who is innocently threatening him, because otherwise Y would kill X (holding no right to do so). The argument has three premises. It endangers Thomson's theory because the first is a theorem to which Thomson is committed: if in a circumstance C, X is (all things considered) justified in killing Y in self-defense, then in C anyone else is (all things considered) justified in killing Y in defense of X. ('The permissibility of X's killing Y in self-defense goes hand in hand with the permissibility of Z's killing Y in defense of X.')¹⁶ As, in our case of the two trains, each side is justified in killing the other in self-defense, it follows that a third party is allowed to kill *either*. (I shall refer to this premise as Transferability.) The other premises of the argument seem self-evident. The second premise is that,

15 See M. Otuska, 'Killing the Innocent in Self-Defense,' *Philosophy and Public Affairs* 23 (1994) 74-94, at 79. Compare Russell Christopher's analysis of such cases under Robinson's theory of justification in his 'Unknowing Justification and the Logical Necessity of the Dadson Principle in Self-Defense,' *Oxford Journal of Legal Studies* 15 (1995) 229-51, at 239-45.

16 Recall that for Thomson, having a right to kill in self-defense is equivalent to being justified in killing in self-defense. In an earlier article, however, Thomson insists that 'it cannot be concluded that you too can do nothing, that you cannot attack [the child] to save your life' (Thomson, 'A Defense of Abortion,' 52). Nancy Ann Davis has argued that against an innocent threat the victim has merely a permission, not a right, to kill in self-defense. The difference is precisely that this does not transfer to third parties. See her 'Abortion and Self-Defense,' *Philosophy and Public Affairs* 14 (1984) 175-207.

in situations in which someone holds a right to self-defense, the right (indirectly) generates an obligation of people who bear a special relation to the right-holder. Specifically, if Z is X's physician, Z ought (all things considered) to take advantage of every right (permission) X has, if this is necessary to save X's life in a 'medical' situation. (Call this premise Obligation*.)¹⁷ The third premise is that (all things considered) moral obligations (as opposed to mere permissions) satisfy the Agglomeration Rule: if, in a circumstance C, Z ought (all things considered) to do *a* and in C, Z ought (all things considered) to do *b*, then in C, Z is obliged to do both *a* and *b*.¹⁸

Suppose that in a medical situation (X and Y are Siamese twins, say) X and Y are innocently threatening each other. Both are Z's patients. According to Thomson's formulation, X is justified in killing Y in self-defense and, hence, by Transferability, Z is justified in killing Y in X's defense. As Z is X's physician, it follows, by Obligation, that Z is obliged to kill Y in X's defense. The case is symmetrical, so the same is true of Y: Y is justified in killing X in self-defense. Hence (by Transferability) Z is justified in killing X in Y's defense. Hence (by Obligation) Z is obliged to do so. It follows, by Agglomeration, that Z is obliged to kill both X and Y, even if they are both innocent; this is, of course, absurd. (Note that the three premises do not imply a paradoxical conclusion if both X and Y are *villainously* attacking each other. After all, in such a situation, neither is allowed to kill the other in self-defense. Hence, there is no way to deduce an all-things-considered obligation of a policeman [say] to kill one in defense of the other.)

As I said, Thomson is committed to the first assumption of this argument; the other two seem intuitively secured. Hence, from Thomson's standpoint, the most plausible response to it points to the required amendment of her formulation: X is justified in killing an innocent threat Y, only if by doing so X would earn a substantial period of life, and *not merely if otherwise Y would kill X*. Hence, in the symmetrical case described above, Transferability implies that Z is justified in killing Y in X's defense

17 As a referee for the *Canadian Journal of Philosophy* pointed out to me, Obligation* is clearly false.

Obligation*: if Z is X's physician, Z ought (all things considered) to take advantage of every right (permission) Z has, if this is necessary to save X's life in a 'medical' situation.

In a certain situation, a physician may be permitted to sacrifice his own life to render medical care for his patient.

18 Hereafter, I shall omit the modifier 'all-things-considered.' But when I say that 'X is justified' I mean that X is all-things-considered justified.

only if X earns a substantial period of life thanks to this action. This is the justification that generates Z's obligation, and if so, Z is obliged to kill Y in X's defense, only if by so doing he does not simultaneously kill X.

III Culpability Does Matter: Against the Weberian Argument

So much for the Intuitive Argument. I shall now show that even if one accepts its conclusion — that is, even if one agrees that the victim has a right to kill the falling man in self-defense — the Weberian Argument must be rejected. There is a deep intuitive difference between the ethics of self-defense against innocent threats and the ethics of self-defense against villainous threats.

The most striking difference has to do with the amendment suggested in the last section. Interestingly enough, in contrast to a victim who defends himself against an innocent threat, the right a potential victim has to kill a villainous aggressor in self-defense seems to be insensitive to how much the victim would gain from this action. To see that this is so, consider a variant of the villainous aggressor case. The aggressor is driving a truck that is heading towards the victim, holding a loaded gun, with which he is about to shoot at the victim. As before, the victim's only chance to stop the aggressor from killing him is by shooting at the aggressor. But note the crucial difference: if the victim would kill the aggressor, the truck would run him over anyway and kill him. Is the victim allowed to shoot at the driver in order to stop the driver from killing him? Most people would answer positively. Note the sharp difference between the innocent and the villainous: had the aggressor Y been morally innocent, we would say that the victim X is *not* allowed to shoot Y if X would die imminently anyway. The culpability of the aggressor is, thus, crucial.

If this is indeed the verdict of commonsense morality, it immediately raises the following question: is X's right to kill a villain aggressor, where X is about to be killed anyway, a right to self-defense? Isn't it a kind of revenge? I am not sure. Suppose X suffers from a terminal disease. It sounds strange to say of X that in fighting back he acts out of vengeance, as if he revenges his own blood. Be that as it may. One thing is clear: however we name the right to fight back against the villain, one possesses this right even if one's death is unavoidable. Hence, one holds a right to kill a villain because of *two* elements, to wit: the right to avoid the harm of death, and the right to fight back, which one holds even if the harmful death is unavoidable. It is, then, possible, that it is the *second* element that makes Thomson's original formulation correct with respect

to the villainous aggressors; you are justified in killing a *villainous* aggressor in order to stop him from killing you. Culpability has a crucial justificatory role. Hence, Thomson is not entitled to derive a right to kill an *innocent* threat in self-defense from the right to kill a villainous aggressor in self-defense. The Weberian Argument, which is designed to do so, must be fallacious.

In fact, Thomson's own attitude towards the Weberian Argument is unclear. She toys with the idea that fault might matter: 'suppose an aggressor will take ... only your left foot unless you kill him. Here the aggressor's fault or lack of fault may well be thought to make a difference.'¹⁹ That is, she admits that there is a possibility that, in some cases, killing in self-defense would not be justified unless the aggressor is a villain. But if the Weberian Argument is sound, it should apply also to the case of an aggressor who is about to take your foot. Specifically, if you were justified in killing the villain in defense of your foot *because he is a villain* — you would be 'dishing out punishment to the villainous for the things that they do.' And it would be permissible to kill him, even if your foot was not in danger. Thus, according to Thomson's own statement, her theory is consistent with what may be called the weak fault-based notion of self-defense:

The weak fault-based notion of self-defense: In some cases in which X is justified in killing Y in self-defense, X is justified in doing so because Y is not free of fault for the particular aggression that Y is committing.

The relevance of the aggressor's culpability is supported by other moral truths: even if killing the aggressor is the safest way to avoid death, nevertheless in case he is innocent, you should take a risk and incapacitate him rather than kill him. There is no such duty if the aggressor is villainous. Further, you have no duty to retreat from a confrontation with a villain who invaded your home, even if you can retreat with complete safety. You should, however, do anything possible to avoid confronting a would be innocent attacker. It thus seems that the Weberian Argument were sound, it would prove more than Thomson wants it to prove.²⁰

19 Thomson, 'Self-Defense,' 289

20 Consider two types of robbers, the bloody-minded and the greedy. If you don't kill the bloody-minded robber, he will kill you unless you let him have your purse. The greedy robber, in contrast, won't kill you. He has no need to take such extreme measures: he is so much stronger than you that he would win a nonlethal struggle and have your money anyway. Shooting at him is the only way to stop him. Should there be a difference between the bloody-minded robber and the greedy one? Some might believe that it is forbidden to kill in self-defense, if you can save your life by

In the next section I shall discuss McMahan's Justice, which seems, at first glance, to be the most natural and straightforward account of the role of culpability in the ethics of self-defense. Justice does not take the Weberian Presumption seriously. Indeed, in light of the obvious falsity of the conclusion of the Weberian Argument, Justice's proponents tend to dismiss the Weberian Presumption against exercising justice-making violence. This is unfortunate: by reflecting on the Weberian Presumption more carefully, a decisive flaw in the justice-based theory of self-defense becomes noticeable.

IV Why Does Culpability Matter I: Against Justice-Based Accounts

As already noted, the Weberian Argument is typically taken to be fallacious. Zohar, for one, says, 'the fact that the context is crucially different from that of a courtroom does not exclude the relevance of the aggressor's guilt.'²¹ Guilt matters not because it justifies a death penalty, but because 'it tips the otherwise balanced scales of 'life versus life.'²² Notably, Zohar seems to agree with Thomson that the violence rightfully exercised by a private individual in self-defense should not be justified on the basis of considerations that justify punishment. The role culpability plays in justifying self-defense is different *in kind* from its role in justifying punishment. To repeat, when culpability justifies killing in self-defense, its only role is to tip the otherwise balanced scales in favor of the potential victim. This formulation suggests a straightforward and natural answer to the question of where the Weberian argument went wrong. In McMahan's Justice it is developed in some detail.

letting the aggressor have (some of) your money. I, among many others, believe that, in the first case, but not in the second, you are allowed to fight back. There are other cases that support this intuition. If a minor attacks you, and you can survive without killing him, at the price of losing a million dollars (you have to throw away a suitcase full of money in order to run fast enough), you must do it. The same is true of a psychotic (innocent) aggressor who attacks you and you can run away. But if a villain attacks you, you need not run away if enough money is at stake. If you didn't flee, you are allowed to kill him in order to save your life. Finally, suppose that you see a very good runner attacked by a psychotic aggressor. The would-be victim prefers not to run away. So he fights back and loses; now his life is in danger. Are you, the observer, allowed to defend him by killing the aggressor? I am not sure at all that a person who could have run and didn't should be saved by killing the aggressor, if the aggressor is innocent.

21 Zohar, 'Collective War and Individualistic Ethics,' 610

22 *Ibid.*, 611

Culpability, according to Justice, makes a person 'liable to harm.' A person is liable to harm only in case 'it is just that he should suffer it, *if* it is necessary that someone must suffer harm.' Hence, Justice 'implies only that culpable action gives rise to liability to harms inflicted in self- or other-defense or self- or other-preservation.'²³ In contrast, punitive violence, in the sense Thomson has in mind, is exercised when a person deserves to be harmed, i.e., other things being equal it is bad if he fails to suffer harm. Put somewhat differently, according to Justice, self-defense is justified in terms of *distributive* justice rather than in terms of *retributive* justice. It is not about inflicting an avoidable harm, but about *shifting* to the aggressor an inevitable harm, which the victim would otherwise suffer; it is not the aggressor's desert that justifies shifting the harm to him, but rather his liability. In other words, killing in self-defense is justified in virtue of the just distribution of harm it brings about.

The proponents of Justice agree that punishment may be meted out only by the government. But they argue that this is not on account of any moral or political principle (defeated in cases of justified self-defense) but rather is primarily a pragmatic matter. We have evolved collective mechanisms for dispensing retributive justice through punishment. Probably, however, private persons would have the right to punish in the state of nature. Similarly, we have collective mechanisms for achieving corrective justice, namely, tort law. But we do not have effective collective mechanisms for dealing with preventive justice, that is, justice in the prevention of harm. This is because in most cases, self-help is the only possible defense in the face of an actual or imminent attack. In short, according to Justice, it is because the state is incapable of administering preventive justice that administering it falls to the hands of private persons.²⁴

Now, proponents of Justice acknowledge that even in circumstances in which self-help is needed, considerations of distributive justice should be restricted. Consider the following statement, made by McMahan:

The harm imposed must, however, be appropriately related to the culpable action... One may not harm a culpable person ... in self-preservation, if his culpability is unrelated to the threat to oneself. For, relative to that threat, the culpable person is a Bystander ... culpability engenders liability only to those harms that are necessary to shift harms caused or made inevitable by the culpable action itself.

23 McMahan, 'Self-Defense and the Problem of the Innocent Attacker,' 260

24 Thanks for Jeff McMahan for this clarification of his position.

This remark is absolutely crucial to the cogency of the distributive-justice conception of the ethics of self-defense. To see why, consider the 'No case' discussed above, in which you stop bullets that are about to hit you by using a bystander who just happens to be around as your shield. Suppose, however, that the bystander is not morally innocent, as he has just murdered someone. Still, intuitively, an intentional killing of a bystander in self-preservation is a murder, even if the bystander is a (past) murderer. Thus, a restriction on the distributive considerations is necessary in order to classify the above culpable bystander case as a 'No case.' Thanks to this restriction, Justice embraces the virtue of Thomson's ethics of self-defense, namely: the blanket prohibition to kill bystanders in self-preservation. Justice is merely more restrictive: it allows defensive killing only if the agent of the potential harm is culpably responsible for it.

Unfortunately, however, as it stands, the restriction formulated in the above-quoted passage is *ad hoc*. In its pure form, Justice does not have the resources for explaining the prohibition on killing a culpable bystander in self-preservation. The reason is simple enough: according to the above-quoted remark, at an earlier time, when the culpable bystander was *about to* commit his murder, he was liable to having an inevitable harm shifted onto him. Now that he has already committed the murder, he is liable no more; he has gained his immunity back, as if he is as innocent as anyone else. But this cannot be warranted by considerations of distributive justice alone: if culpability is all that matters with respect to the distribution of inevitable harms, then why doesn't the wicked bystander's culpability tip the otherwise balanced scales of 'life versus life'? How does the fact that a person has already harmed someone remove his liability to harm? Indeed, it seems that, in terms of distributive justice, shifting harm to a culpable bystander might be fairer than shifting harm to a culpable aggressor. After all, the culpable bystander has already committed a crime, while the culpable aggressor has as yet harmed no one. A 'pragmatic' explanation for the prohibition to kill the culpable bystander, in terms of private versus collective action, won't do; for in the circumstances under discussion, self-help is the only way to distribute the inevitable harm justly. In short, restricting liability to harms to which the culpable action is 'appropriately related' is crucial. On the conception of self-defense provided by Justice, however, it is philosophically ungrounded.

Of course, it might be argued, on behalf of Justice, that acting in self-defense is just, because unless the victim kills the villain, the villain would kill the victim. In contrast, the culpable bystander has done nothing to threaten the victim's life — his culpable behavior is *causally unrelated* to the inevitable harm the victim is about to suffer. Hence, it is unjust to kill him in self-preservation. It might be thought, thus, that the

traditional justice-based accounts can avoid my objection by insisting that liability to certain types of harm is rigidly connected to culpability for certain types of action. In no sense is general culpability sufficient to generate liability to harm.

Now, I do not deny that this is how justice-based accounts are usually structured: they take causal relation to a specific unavoidable harm to be necessary for the *justice* of defensive violence aiming at the prevention of this harm. I argue, however, that this dual structure of justice-based accounts is unstable. In particular, as it commonly interpreted, Justice is inconsistent with *any* appeal to causal relations. This can be seen through the critique by Justice's proponents against Thomson's analysis of innocent threat cases. Thomson appeals to mere causal relations, and such a move, they claim, nullifies the moral character of the right to self-defense. Causal relations, proponents of Justice insist, do not constitute a morally relevant difference between innocent threats and innocent bystanders. My question is this: if so, how can causal relations contribute so dramatically to the justice of killing in self-defense in culpable aggressor cases? If causal relations are morally relevant to the extent that they distinguish between culpable bystanders (the killing of whom in self-preservation is prohibited) and culpable aggressors (the killing of whom in self-preservation is permissible), why are causal relations irrelevant to distinguishing between innocent threats and innocent bystanders?

Put somewhat differently, Justice's failure in explaining our intuitive responses to the culpable-bystander type of case is parallel to the failure attributed by Justice's proponents to Thomson's ethics of self-defense to explain our intuitive responses to the innocent bystander case. It is claimed, against Thomson, that the distinction between innocent threat and innocent bystander is morally arbitrary. If so, I have just argued, Justice's distinction between culpable aggressors and culpable bystanders is equally arbitrary. True, the roles of causation in Justice and in the Thomsonian rights-based theories of self-defense are different — in the former theory, causation is a necessary condition for liability to defensive violence, whereas in the latter, causation is sufficient. Yet, proponents of Justice reject Thomson's rights-based account because, in their view, causal relations are morally *insignificant*. Unfortunately, this is inconsistent with making causal relation morally necessary for justice in the prevention of harms.²⁵

25 McMahan tends to bite the bullet. See J. McMahan, 'Self-Defense and Culpability,' forthcoming in *Law and Philosophy*; and 'The Basis of Moral Liability to Defensive Killing' *Philosophical Issues* 15 (2005) 386-405.

There is one attractive solution to the problem of the culpable bystander that can be embraced by Justice.²⁶ I believe, though, that we should resist this way of solving the problem, as well. The solution I have in mind involves a partial acceptance of the Weberian Presumption (and hence leads to an important modification of Justice — I shall call the resulting theory ‘Weberian Justice’ or ‘W-Justice’). According to this solution, the right to self-defense should be explained morally and ‘politically.’ To be specific, an ethics of self-defense has to explain how, in a case where killing in self-defense is justified, the strong presumption against shifting harms is defeated. (W-Justice’s answer is the one given by Justice, in terms of liability.) In addition, it has to explain how the Weberian presumption against private violence is defeated; why is a *private* person allowed to shift harms? Now, at this point only, causal relations *are* relevant: if a person is committing aggression against *you*, you are allowed to fight back. Only in such a circumstance is the presumption against private violence defeated (and since it was defeated, any other private person is allowed to kill the aggressor in the victim’s defense).

Now, according to W-Justice, Thomson fails adequately to defend the distinction between innocent bystanders and innocent threats because the problem she faces is moral rather than political. The question she has to answer is why shifting harm to the innocent threat is allowed. With respect to this question, causal relations are useless; the presumption against shifting harm is defeated only where the aggressor is culpable.

I have three points to make in relation to this modification of Justice. Admittedly, they do not add up to a conclusive argument against W-Justice; they do, nonetheless, cast doubt on it. First, W-Justice is faced with a difficult question. According to W-Justice, shifting the harm to the culpable bystander is just. So why is this by itself not sufficient for defeating the presumption against private violence, in cases the victim would survive by self-help? Second, W-Justice is faced by an obvious

26 There are other solutions that seem to me unattractive. According to one of them, self-defense is morally justified, because the aggressor is culpable *and* his culpable aggression is causally relevant to the inevitable harm to be shifted to him. According to this solution, a mere causal relation makes killing the innocent aggressor in self-defense *more* justified, but unless it is accompanied by culpability, this is an insufficient justification for shifting harm. Proponents of Justice tend to reject this explanation. They tend to believe that there is *no morally relevant* difference between an innocent threat and an innocent bystander. Alternatively, it might be claimed that mere causal relation — the fact that one is part of a threat which is to cause an inevitable harm — has *no* moral significance unless one *culpably* constitutes a threat. But this has to be explained. And the only explanation I see is presented in the text.

counterexample: it implies that the Weberian Presumption against private violence is defeated in circumstances in which, in fact, it is not. This is because W-Justice implies that, in some circumstances, blood revenge is permissible. But this practice is what modern states resist by monopolizing the right to exercise violence. Here is the counterexample. Suppose that Y murdered X's wife. For some reason, Y was acquitted in a trial. X's life would be ruined unless he revenges the blood of his wife. So by revenging his wife's blood, X is shifting from himself to Y an otherwise inevitable harm to which Y is culpably related. You might think that killing Y is not proportional to the harm X has to suffer. But this, of course, does not have to be the case; X won't be able to rebuild his life if he does not see justice done.

This last observation leads to the third point. According to W-Justice, the presumption against shifting harms and the presumption against private violence are unrelated: it is possible that shifting an inevitable harm to a culpable bystander would be *just* or *fair*, yet prohibited because the Weberian Presumption was not defeated. But, as is strongly suggested by the above observations, the separation between the moral and the political is shaky. I thus conjecture that the Weberian Presumption is stronger: just institutions are essential to, or constitutive of, distributive justice in the sense that *violent* distribution is not just unless an impartial institution carries it out. True, a villainous aggressor is liable to harm by virtue of being culpable. But you, as a private person, are not allowed to violently shift harm in order to bring about just distribution of harms. This is not because of an external pragmatic restriction, but because doing so won't bring justice. Justice (of violent distribution) is a property of acts of institutions rather than of private individuals.

Here is an example that supports this general conjecture.²⁷ Suppose you believe in equality — in your opinion it is unfair that some people, through no fault of their own, are worse off than others. You live in a society whose basic structure is so designed that it optimally promotes the ideal of equality. Unfortunately, however, there are private companies — like the bank that recently hired you — that could do more towards a fair distribution of resources. So in order to further promote equality, you rob the bank, and transfer some of its money to the poor. How much better, with respect to fairness, is your society now than it

27 I borrowed this example from the Pogge/Cohen debate about Cohen's incentive objection to Rawls's theory of justice. See T. Pogge, 'On the Site of Distributive Justice: Reflections on Cohen and Murphy,' *Philosophy and Public Affairs* 29 (2000) 136-69, at 162. Cf. G.A. Cohen, 'Where the Action Is: On the Site of Distributive Justice,' *Philosophy and Public Affairs* 26 (1997) 1-27.

was before your act of violent robbery? The answer might well be *no better at all*. True, the distributions of benefits, burdens, and harms are descriptively more equal. Yet the outcome fails to be fairer, because of the *way* it was generated. This suggests that the justice of violent redistribution is *essentially* institutional. If so, liability, as McMahan defines it, is *irrelevant* to the justification of the private violence exercised in self-defense.

Two implications of the strengthened Weberian Presumption against private violence are noteworthy. First, the strengthened Presumption denies that enforcement of justice (of any type, retributive, corrective or preventive) can be administered in the state of nature — *if it involves violent harming*. That is, in the absence of a sovereign, killing a murderer is revenge rather than punishment (whatever the intentions of the killer are). In the same spirit, the right to kill in self-defense cannot be explained on the basis of considerations of distributive justice. Note that the strengthened Presumption has to do with *violence* more than it has to do with punishment and retribution. Secondly, the strengthened Presumption asserts that a legal system in which vigilantism is permitted, or prescribed, would be unjust — the state should not authorize private people or companies to do the ‘dirty’ work for it. Only an impartial just institution can justly exercise legitimate violence. (Compare the strengthened Presumption to the Rawlsian belief that private charity cannot *replace* social justice. Sovereignty is essential for the justice of redistribution; in the absence of a sovereign, private good willed individuals could not achieve it.) This might explain the common belief that it matters how and by whom violent punishments are administered, and that centralizing, or collectivizing, the coercive justice-making system is necessary for establishing a just society.

The strengthened Weberian Presumption is thus inconsistent with W-Justice, for it denies the relevance of the aggressor’s liability to the justification of private self-defense. True, according to the strengthened Presumption, culpability makes a person liable to an inevitable harm. By itself, however, the liability of a culpable aggressor cannot explain why a private person holds a right to kill him. If only just institutions can violently force fair distribution, W-Justice is wrong: a right to self-defense cannot be justified on the ground that the potential victim is doing justice by shifting a harm that he is about to suffer. As I said, I did not complete the case for the strengthened Presumption: the above points do not add up to a conclusive argument against W-Justice. They constitute, however, a powerful reason to look for another explanation of the role of culpability in the ethics of self-defense.

V Why Does Culpability Matter II: An Interest Theory of the Right to Self-Defense

So why does culpability matter? I shall advance, in this section, a Thomsonian explanation of the right to kill an innocent threat in self-defense, and a new account for why the culpability of the aggressor is relevant to the permission to shift harms to him. My explanation amounts to an amendment (rather than a rejection) of the Thomsonian ethics of self-defense. Let me begin by reformulating three Thomsonian convictions that my account fully embraces.

First, properly modified, the conclusion of the Intuitive Argument presented in section II is correct. The permission to kill an innocent threat Y in self-defense follows from the brute fact that, unluckily for Y, Y violated X's (Hohfeldian) claim against Y that Y not threaten X's most basic interest (staying alive). In light of the discussion in section II, it is clear that the permission to kill in self-defense depends on whether X has an interest in killing Y because of this violation. And this strongly suggests an interest theory of the right to self-defense: Y's violation of X's claim against Y generates a permission (a Hohfeldian liberty) to kill Y, *only if it is in X's interest to kill Y*. Thus, the interest theory of rights neatly explains why one is not allowed to kill an innocent threat in self-defense if one gains nothing substantial thereby. On such a theory, a person holds a right by virtue of one of his interests. In the absence of such an interest, he has no right of self-defense. The interest theory has another desirable feature. It *might* explain why the *numbers* of the innocent threats count, that is, why it seems much more problematic for one person to kill eight innocent threats in self-defense than to kill only one. The proportionality constraint with respect to numbers is indeterminate, as we have seen in section II. I shall have more to say on this indeterminacy later.

Secondly, innocent threats are distinguished from innocent and culpable bystanders because the latter do not violate any claim the victim has against them. This raises Justice's objection regarding the moral difference between innocent threats and innocent bystanders. Yet, we saw that both Justice and W-Justice are open to a similar objection regarding the moral difference between culpable bystanders and culpable aggressors. True, W-Justice uses this distinction for addressing a 'political question' — but I argued, I hope persuasively, that the very distinction between the moral and the political is unstable. Finally, Thomson's appeal to the Hohfeldian conceptual apparatus further clarifies the distinction between bystanders and threats. The right to kill Y in defending X is related to the fact that Y (rather than anyone else) violated a claim X has against Y. Hohfeldian claims, duties, and privileges are essentially relational.

Let us turn to the asymmetry between innocent and culpable threats. How is the difference that the aggressor's culpability makes to the ethics of self-defense to be explained? This question is particularly difficult in light of the lesson drawn from section III: intuitively, if Y is committing an act of villainous aggression against X, X is allowed to kill Y, even if by doing so, X will not avoid imminent death. Thus, at a first blush, it seems that explaining the scope of the permission to fight back would employ considerations of desert and retribution: the culpable aggressor is liable to defensive killing by virtue of the punishment that he deserves. And this, of course, is inconsistent with the Weberian conviction on which Thomson bases her conception of self-defense.²⁸

I shall argue that this inference to inconsistency is too hasty: despite appearances to the contrary, the right to 'pointlessly' fight back can be explained without rejecting the strengthened Weberian Presumption. For, there is a special interest that a victim has to stop a villain from killing him. And, if such an interest exists, the victim's right to fight back is preventive rather than retributive. Killing a villainous aggressor in self-defense is permissible because, in doing so, the victim defends a special kind of self-interest. To show that there is such an interest, let us try to address a preliminary question, namely: why killing is wrong. The answer is surprisingly complicated. To see one particularly relevant complication, let us take a quick look at a puzzle McMahan recently formulated in a completely different context.

A runaway trolley is careering down the mainline track. If it continues along this track, it will crash into the station, killing hundreds of people.... it can be diverted onto one or the other of two branchline tracks. You are a bystander who happens to have access to the switch that can divert the trolley. You see that there is an innocent bystander on each of the branchline tracks: Young is on the track to the left, and Old is on the ... right ... neither will be able to get out of the way of the trolley if you divert it.²⁹

Suppose that you say, as so many of us would, 'OK, if all other things are equal, I shall divert it to the left branchline, where Old is, if he is older than Young.' The question McMahan is asking is: if so — if killing Old is preferable to killing Young — is it also the case that murdering an old person is less wrong than murdering a young person? Should we infer from our intuition about the above version of the Trolley problem that

28 Y's liability won't do, for, in the case under discussion, in killing Y, X does *not* shift an inevitable harm to Y, because by killing Y, X does not avoid any harm.

29 Jeff McMahan, *The Ethics of Killing* (Oxford: Oxford University Press 2002), 237

if Raskolnikov preferred murdering an old rich lady to murdering her daughter, his act of murder was less wrong than it would have been had he murdered the young woman? 'The common view ... is that the wrongness of killing persons does not vary with such factors as the degree of harm caused to the victim, the age, [etc.]... Our understanding of the wrongness of killing should reflect our belief in the fundamental moral equality of persons.'³⁰

This observation should lead us to distinguish two senses in which life is valuable. In one sense, the character and the contents of a life determine the value of that life *to* the person who lives it. But in another sense, the value of life is the value *of* the person who lives it. In this latter sense, the value of life is the same thing as the moral worth of a person. It is determined by the nature of the subject of this life, by the capacities that make that individual the kind of being that he or she is.³¹

The two types of value provide different reasons for the moral prohibition of killing. In the typical case, killing is harmful. Yet, the extent to which killing is harmful varies. The reason why every murder is equally wrong is, thus, rooted elsewhere. Aside from being harmful, murder is wrong because it involves a failure of respect for the worth of the victim. McMahan concludes that the morality of killing is 'two-tiered.' It is partly a set of obligations generated by material interests; these obligations focus on the effects our actions have on the interests of others. The other tier is the morality of respect, which is made up of constraints on our behavior 'that spring from our recognition of others as mature agents on an equal moral footing with ourselves.'³² While the badness of death is correlative with the value of the victim's possible future life, the wrongness of killing is correlative with the value or worth of the victim himself.³³

The two-tiered morality of killing naturally leads to a two-tiered ethics of self-defense (a version of which I shall develop shortly). I shall argue that when the victim defends himself against a villainous threat, he is trying to avoid both the harm of death and also disrespectful treatment at the hands of the aggressor. The latter involves the negation of the value *of* the person.

30 *Ibid.*, 235

31 *Ibid.*, 241

32 *Ibid.*, 246.

33 *Ibid.*, 242

Before further developing this idea, I should like to clarify the notion of respect that I shall exploit. Indeed, there is no philosophical consensus regarding the morality of respect. The approaches defended in the literature can be crudely reduced to two. The common one, and supposedly the one McMahan has in mind, might be called 'Kantian'; it inspires the infamous conception of 'rights as trumps.' The two-tiered ethics of self-defense that I shall develop exploits a *non*-Kantian conception of respect and (consequently) of rights.

In the Kantian view, respect has two important features: the 'loci' of disrespect is intentions and actions rather than outcomes — i.e., a murder is a disrespectful *killing* which manifests the murderer's disrespectful *attitude* towards the victim; if at all, the outcome — the death itself — might be disrespectful only in a derivative sense. Secondly — and more importantly — the disrespectful treatment does not, as such, inflict harm on the victim. Suppose it is inevitable that I will be killed but I can choose between two ways of being killed. One way, the killing will not be disrespectful, but it will result in immediate death. The other way, I will be killed disrespectfully; I will not, however, die immediately, but will live for another week. In the Kantian view, my reason for avoiding the disrespectful killing has nothing to do with my interests. That is, it is in my interest to be killed disrespectfully if I would thereby live longer than I would if I were killed in a way that did not involve disrespect. Now, rights-based moralities tend to rely on this notion of respect. Hence, the conception of rights as trumping utility: 'rights ... are based neither on right-holder's interests nor on that of others. Rather they express right-holders' status as persons and the respect owed to them in recognition of that fact.'³⁴

Contrary to this view, I believe that we have an *interest* in being treated respectfully. And, in particular, we have an interest in choosing the way we die. Consider again the above story, somewhat modified: a person can either die immediately, or have another week of good life followed by a long period of coma in which he will be totally dependent on others. This mode of existence seems to the person to be disrespectful. His interest in a respectful death might explain his understandable decision to die immediately. Thus I claim that the morality of respect includes a morality of interest of a special kind — a moral, nonmaterial interest that we have by virtue of being persons. One consequence of being a person — of having properties that place one above the threshold for having

34 Raz, *The Morality of Freedom*, 188. The conception of rights as 'trumps' is Dworkin's. See his 'Taking Rights seriously,' in *Taking Rights Seriously* (Cambridge, MA: Harvard University Press 1977).

personal moral worth — is having, in addition to material interests, moral interests that generate an additional layer of moral rights.

I cannot provide here a full defense of the non-Kantian conception of respect. Nor can I defend the theory of rights that it inspires. Instead, I shall shortly present Raz's critique of the 'rights trumps utility' conception of rights. Assertions about rights, Raz stresses, are true by virtue of the interests of the right holder — X has a right if an aspect of X's well-being is a sufficient reason for holding some other persons to be under duty toward X³⁵ — hence, 'it cannot be said that [rights] are trumps in the sense of overriding other considerations based on individual interests.'³⁶ The right to respectful treatment is no exception. In the broadest sense, one's right to respectful treatment is one's right to have one's interests taken into account by others. More strictly interpreted, the right to respectful treatment is deeply related to our interest in deciding freely the course of our lives. Respecting people 'as people' consists in giving due weight to their interest in having and exercising that capacity. Respect, in its stricter sense, has to do with *this particular interest*.³⁷ Murder is wrong, not just because of its being materially harmful, but also because the murderer does not give due weight to the victim's interests. In the murderer's eyes, the victim's most basic interests simply do not count; this attitude is the most extreme form of disrespect.

I am now in a position to present the non-Kantian version of the two-tiered ethics of self-defense I favor. The right to kill an *innocent* threat in self-defense is the right to avoid the badness of death — it is a right to act in one's material self-interest, against those who violate their duties not to hurt him. In such cases, there are no moral interests involved; the falling man does not treat the victim disrespectfully. In contrast, the right to kill the villain has a broader basis; the victim has (also) a right not to be wronged by disrespectful treatment, a right that the culpable aggressor violated. The culpable aggressor manifests disrespect of the victim by giving no weight to this interest in respectful treatment. He has his right not to be killed partly because of this.

This move explains why the relevance of culpability to the ethics of self-defense is consistent with the strengthened Weberian Presumption. For, in the view I have just outlined, the permission to kill a villain has nothing to do with the fact that the villain is *liable* to harm. Quite to the

35 Raz, *The Morality of Freedom*, 166

36 *Ibid.*, 187

37 *Ibid.*, 190

contrary: distributive considerations have nothing to do with the moral permission to kill in self-defense. The right to kill culpable aggressors in self-defense springs (partly) from the moral right to avoid disrespectful death, which, in turn, springs from the victim's interest in respectful treatment. A person has such an interest even where his death is unavoidable. This approach yields a wholly different understanding of where Thomson's Weberian argument went wrong. According to Justice and W-Justice, Thomson confuses desert and liability by mistaking retributive justice for distributive justice. According to the explanation advanced here, she fails to see the scope of interests defended by the right of self-defense.

1. An Objection from the Case of a Non-Threatening Culpable Aggressor.

Here is a worry concerning the two-tiered ethics of self-defense. Suppose someone deliberately and disrespectfully intends to negate your worth by killing you, but his attempt will fail because his gun is not loaded — and you know this in advance. The aggressor is, therefore, 'a *non-threatening* culpable aggressor.' It seems that, according to the two-tiered ethics of self-defense I recommend, you are entitled to kill the would be attacker anyway in order to prevent him from attacking and thus committing an act of disrespect against you; for his attempt to murder you is as disrespectful as actual murder. This, however, sounds absurd. To put the objection differently, if an unsuccessful villainous attempt to murder X is as disrespectful to X as a successful attempt would be, then preventing the murder by defensively killing the villain does not prevent the disrespectful treatment; for, the victim was already treated disrespectfully by the attempt. Hence in the absence of a material interest to avoid being killed by the aggressor, killing him is not an interest of the victim. The objector concludes that 'pointless' self-defensive killing can be only an individual's exercise of punitive violence, hence a paradigmatic transgression of the Weberian Presumption.

It should be noted that McMahan's explanation of the equal wrongness of all murders is subject to a similar worry. If the disrespect involved in murders makes them all equally wrong, why doesn't the disrespect involved in failed attempts make them as wrong as murders?³⁸ Be that as it may, it seems that the solution to the problem is as

38 Some philosophers would bite this bullet — they would claim that morally speaking, attempts are as wrong as murders — this is, however, counterintuitive.

follows: however wrong failed attempts are, we have no right to compel people to have respectful attitudes towards us. We have a right only to prevent people from *harming* us in ways that are disrespectful. In other words, it seems that considerations of respect provide *only* reasons that make it wrong to kill us, and so give us reasons to prevent ourselves from being killed. These reasons do not entitle us to compel people to *feel* respect for us, or to refrain from unharmed acts of disrespect, or to acknowledge our worth in other ways. The question is, of course, *why* is this so.

I can see two reasons. In order to present the first, I have to point to another ingredient in Raz's interest theory of rights. In some cases, Raz observes, the right-holder's interest itself, conceived independently, is deemed insufficient to justify holding others to be subject to the extensive duties and disabilities commonly derived from having a right. Raz claims that in the Common Law, freedom of expression is such a right: it is regularly defended not merely on the grounds of the individual interest in free speech but rather on grounds of *the public interest*.³⁹

So let us suppose, for the sake of the argument, that in a case of villainous aggression, the harm inflicted on the victim by virtue of the disrespectful treatment is the same whether or not the victim is killed as a result of the aggression directed against him. By itself, the victim's interest in preventing the disrespectful treatment is insufficient for generating a right of self-defense. Still, the interest of the public is relevant. The culpable aggressor threatens to violate not only the respect due to the victim, but also the social order: any successful crime damages public security much more than unsuccessful attempt. Killing him before he murders is thus a prevention of a substantive public harm. If this is true, we have a Weberian justification for killing a culpable aggressor even where the victim will die soon anyway.⁴⁰ There is no parallel justification for killing the falling man or non-threatening aggressors — in these cases they commit no crime; these individuals inflict almost no harm on the public.

But there might be a deeper reason for the prohibition on killing a non-threatening culpable aggressor. For, perhaps, the disrespect involved in a failed attempt is not as crucial to the victim's well-being as disrespectful death. It is, after all, a fact of life that people care a lot about how the last moments of their life would look like. They have strong preferences about the way they die. In contrast, instantaneous disre-

39 Raz, *The Morality of Freedom*, 179

40 I thank Noam Zohar for this suggestion.

spectful treatments that a person can overcome are usually isolated and forgotten. Of course, not every preference of ours reflects a genuine interest and not every genuine interest is reflected in our preferences. Still, preferences are good evidence for the existence of such interests.

Thus, a possible response to the objection from the case of a non-threatening culpable aggressor would come down to this. In general, the harm prevented by self-defense against culpable aggressors is greater than the harm prevented by stopping a non-threatening aggressor from committing a failed attempt even if an imminent death of the victim is unavoidable. Hence we have a right to kill in order to prevent a disrespectful death, but we have no right to kill in order to prevent a disrespectful non-threatening attack. (Note that this is *not* a Kantian response: in the Kantian view, the response should be 'respect-based' *rather than* 'harm-based'. But remember that within the non-Kantian framework I have adopted, the morality of respect includes a morality of moral [rather than material] interests.)

2. *An Objection from Proportionality.*

There is another serious worry that the two tiered ethics of self-defense should address. Consider the pointless defensive killing that the two-tiered ethics allows. Is the interest of avoiding disrespectful death a strong enough interest to justify such permission? Furthermore, are the reasons involved in respect strong enough to adequately explain why one may kill a *very large* number of malevolent aggressors to save his life (in light of the fact that one cannot do the same in the case of innocent threats)? The answer to these questions seems to be no. But if so, the two-tiered conception of self-defense is inadequate because it entails too weak a proportionality constraint.

Three different reactions to this worry might be suggested. It might be argued that the objector misses a factor that makes killing in the above problematic cases proportional. Alternatively, one might argue that the killing in these cases is *disproportional*, and that, because of some special feature of culpable aggressor cases, the victim has a *right to act disproportionately*. According to the third response, the very concept of proportionality exploited by the objector is misguided. For, the objector's concept is borrowed from the 'realm of distributive justice.' But the realm of justice is unrelated to the 'realm of rights' — and the proportionality that constrains the right of self-defense has nothing to do with rules of justice. In the next section I shall develop this third line of argument. Before doing so, I shall make some brief suggestions about developing the other two.

First, it might be claimed that the interest in respectful death generates a right because of the public interest. Hence, the calculation of the proportionality by which the victim's right of self-defense is constrained should take this other interest into account: even if the victim kills an aggressor and then dies, he prevents a harm that would be otherwise inflicted on the public. Thus the response in these cases may be proportional after all.

The second reaction would concede that the right to kill culpable aggressor(s) in self-defense is subject to the proportionality constraint only to a very limited extent. Yet, limited proportionality is all that is needed in cases that involve *culpable* aggression. To see why, consider the following legal fact: while being engaged in criminal activity, a thief has no right not to be attacked in case it is *unclear* whether or not he poses a lethal threat. (The thief holds such a right only if it is perfectly clear that he poses no lethal threat). That is, the thief's right not to be attacked by disproportional force is dramatically diminished because he is engaged in criminal activity. Why? The reason seems to be this: when a thief violently breaks the law, he *gives up* his right to be treated in proportion to the threat he *actually* poses. Likewise, culpable aggressors have no right to be treated in proportion to the harm they are about to cause — because, by engaging in violent activity, they waive that right.⁴¹

Here is the third reaction to the proportionality challenge. In terms of *distributive* justice, it is indeed *unjust* to pointlessly kill a culpable aggressor in self-defense. For one thing, such a killing is a violation of the 'lesser evil' principle. Clearly, the aggressor does not deserve the death penalty (especially in light of the fact that he has murdered no one.) So, if the victim is about to die anyway, killing the aggressor in defense of the victim brings about more evil rather than less. By killing the aggressor the victim gains a small benefit — he avoids disrespectful death — but the evil the aggressor suffers outweighs this good. However, in general, the proportionality constraint that governs the right of self-defense is *not* a rule of distributive justice. And, in particular, the right of self-defense is not directly constrained by the principle of lesser evil. This becomes clear, I shall argue in the next section, given the fact that innocent threat cases are Yes cases, and given the indeterminacy of the proportionality constraint with respect to numbers. Hence, pointless defensive killing might be proportional — in accordance with the standards of proportionality in the realm of rights.

41 I develop this line of argument in my 'A Defense of the 'Traditional' War-Convention.'

In the next section, I shall explore the very possibility of a right to do what is unjustified. I shall then show that, the right of self-defense should be understood, in some of the Yes cases, as a right of this type.

VI A Right to Do Wrong?

The two-tiered ethics of self-defense, as it is developed here, has an important advantage over alternative theories. Suppose, first, that Thomson is right in claiming that you are permitted to kill an *innocent* threat in self-defense. Second, suppose that, if necessary for your survival, you hold a right to kill two or three *innocent* threats in self-defense (but not eight, say). Suppose, finally, that you have a right to kill a culpable aggressor even if you know that you will die immediately anyway. The two-tiered ethics of self-defense would leave the following question open: Should you take advantage of your rights in these cases? Is it morally justified to use this permission, or rather, is the permission (to kill an innocent threat, say) a *right to do (what is all-things-considered) wrong*? Note that Thomson's Intuitive Argument does not leave room for such a possibility: for her, having a right to kill in self-defense implies being justified in doing so. I shall argue that the excuse/justification distinction is too crude. A plausible version of the two-tiered ethics of self-defense entails that in the typical case, defensive killing is morally justified, but that in some cases, killing an innocent threat, though morally permitted rather than merely *excused*, is, all things considered, morally unjustified.⁴²

I shall do so in two stages. First, I shall show that, in general, recent right-based ethical theories allow for the seemingly paradoxical phenomenon of a right to do wrong. Second, I shall argue that a plausible version of the two-tiered ethics of self-defense implies that the right to kill innocent threats is a right to do wrong.

Let me then dispel the false impression that the very concept of a right to do wrong is paradoxical. Consider the theory of natural rights associated with Robert Nozick.⁴³ In his view, one is entitled to the goods one

42 See Jeremy Waldron, 'A Right to Do Wrong,' in *Liberal Rights* (Cambridge: Cambridge University Press 1993), esp. 66, 77; cf. W.G. Galston, 'On the Alleged Right to Do Wrong: A Response to Waldron,' *Ethics* 93 (1983), 320-4; Waldron, 'Galston on Rights,' *Ethics* 93 (1983), 325-7; and D. Enoch, 'A Right to Violate One's Duty,' *Law and Philosophy* 21 (2002), 361. Cf. Judith Thomson, *The Realm of Rights* (Cambridge, MA: Harvard University Press 1990), 48.

43 See R. Nozick, *Anarchy, State and Utopia* (New York: Harper and Row 1974), ix.

properly (i.e., without violating others' rights) appropriated, in the sense that under no conditions is an individual or institution allowed to coercively expropriate these goods. Thus, suppose you own one pill that can save five young people, but that a rich old man, who needs the entire pill to survive, offers you much more money than the five can pay. In the view under discussion, you are entitled to this pill, so you have an inviolable right to sell the pill to the rich man. This is a moral right, rather than a legal or political right. Indeed, within Nozick's framework, the fact that entitlements are inviolable is one of the most basic *moral* truths. And still, Nozick might believe that, notwithstanding your right to do with the pill whatever you like, you *should* forego the money in order to save as many lives as possible. This is, as far as I can see, a perfectly consistent position. It follows that there might be a moral right to do what is all-things-considered unjustified.

A similar example can be found in Thomson's defense of abortion. Admirers of a famous violinist kidnapped you, and plugged the violinist's circulation system into yours. A kidney ailment from which the violinist suffers made this procedure crucial for his survival; your kidneys are now used to extract poisons from his blood. All you have to do in order to save the violinist's life is to stay connected to him for another five minutes. Thomson seems to believe that your moral right to your body is so basic that you have a right to disconnect yourself whenever you like.⁴⁴ Still, the claim that you should, from a moral point of view, give up your right, seems perfectly consistent with her position. If so, it follows that you possess a right to do what is all-things-considered wrong, or (as Thomson would put it) indecent. Again, I can see no conceptual instability in such a view.⁴⁵

Finally, consider a famous story told by Williams in his critique of utilitarianism.⁴⁶ Suppose that Jim is confronted with the following choice: he can actively kill one of Pedro's (innocent) prisoners, or else Pedro will kill all twenty of them. Jim chooses, egoistically, to keep his

44 See Thomson, 'In Defense of Abortion,' 63, where she discusses the Minimally Decent Samaritan.

45 The terminology I use is not Thomson's. She says it would be *indecent* not to allow one's body to be used for rescue, not that it would be *wrong*. She seems to believe that decent people are worthy of praise if they waive their rights. This appears to be quite close to the notion of supererogation.

46 The example is borrowed from B. Williams, 'A Critique of Utilitarianism,' in *Utilitarianism; For and Against*, J.J. Smart and B. Williams, eds. (Cambridge: Cambridge University Press 1973).

hands clean — he does not want to actively kill one of the innocent prisoners. I think that a helpful conceptualization of Williams's intuition would be as follows: Jim has the moral right to make such a choice, even if this is a right to act wrongly. This right is a right to violate a moral obligation because the numbers morally count; Jim's obligation is to do everything to save the twenty.⁴⁷

Now, the possibility of a right to do wrong is naturally embedded in the dualistic structure of the two-tiered morality in general, and in a two-tiered ethics of self-defense in particular. To see that this is so, note that the Nozick-like or Thomson-like right-based ethical theories I have sketched above are similar in one respect: they locate the realm of rights within a broader moral realm, while insisting that these realms do *not* overlap completely. The two-tiered morality of killing sketched in the previous section has a similar dualistic structure. Furthermore, according to the interest theory of rights I use in constructing the two-tiered ethics of self-defense, X holds rights by virtue of some of his interests. Hence, it is natural to identify the morality of interest and the realm of rights. After all, X's interests are the ultimate basis of the (Hohfeldian) claims X has against others, and of the (Hohfeldian) liberties to act in his interests by harming those who violate their (Hohfeldian) duties to him.

The identification I propose — that of the realm of rights with the morality of interests — substantiates the lesson of my discussion of rights-based ethical theories: in some circumstances, a person ought to act in the interests of others by forgoing his own *rights*. This obligation springs from the morality of respect, i.e., 'from our recognition of others as mature agents on an equal moral footing with ourselves.' Respecting people involves treating them impartially: giving due weight to their interests, even if one is allowed (holds a right) to ignore others' interests and act in one's own self-interests.⁴⁸ More specifically, I shall say that a person is justified in acting in a certain way, if an omniscient, impartial, benevolent spectator would like him to act in this way. The preferences of such an impartial spectator (in the tragic circumstances that are here under discussion) presumably are sensitive to at least three factors. First, she prefers an act that minimizes the magnitude of an inevitable harm. Second, she prefers that the minimum possible number of people be

47 Williams himself never explicitly said that Jim has a right not to kill anyone and thereby let Pedro kill twenty.

48 See Raz, *The Morality of Freedom*, 188; Harry Frankfurt, 'Equality and Respect,' reprinted in his *Necessity, Volition, and Love* (Cambridge: Cambridge University Press 1999), 152-4.

harmed. Third, her preferences are sensitive to considerations of desert and fairness. The impartial spectator would prefer an inevitable harm to be shifted to the person who deserves it, or, to the person who most deserves it, compared to the relevant others.

To apply this abstract idea, let us focus on the story of the violinist. The interest of the violinist to stay alive does not generate a claim he holds against you that you will help him to survive, so in disconnecting yourself, you violate no Hohfeldian claim that the violinist holds against you. Yet, respecting him involves impartially weighing his interests. Staying connected to the violinist is time-consuming. But in doing so, you save the violinist's *life*. So a failure to leave your circulation system plugged into his for another five minutes constitutes a deep failure of respect of the violinist's worth.⁴⁹

Conclusion: in a plausible elaboration of the two-tiered ethics of self-defense, a right to kill in self-defense does not *automatically* imply a *moral justification* to kill in self-defense. That is, one is morally justified in killing in self-defense only if doing so is justified by considerations available to the benevolent spectator (whose preferences are determined by an impartial assessment of the interests at stake). Indeed, I shall shortly argue that, on this conception of justification, the right to kill an innocent threat in self-defense is a right to do what is, all things considered, unjustified.

Before doing so, a further feature of the morality of respect should be noted. The requirement of impartial weighing of interests leads immediately to the conceptual possibility of symmetrical cases. Consider Sophie's forced choice between one of her two children. She is forced to determine who would be killed in a Nazi concentration camp. Impartiality requires dismissing considerations that regard the beauty of one of her children, the fact that she loves the one more than the other, etc. This type of consideration is irrelevant to the issue at hand, hence, justifying any choice by such considerations would be disrespectful to one of the children (or to both of them). Now, in other circumstances of symmetrical choices — e.g., when one has to choose between two equally good apples — the rational strategy would be *to pick* (arbitrarily) rather than *to choose*. But picking one of one's own children is morally wrong as well. Some philosophers infer that one would act (all-things-considered)

49 Thomson would probably say that I have made my impartial spectator too utilitarian. In her version, the impartial spectator *would* prefer the death of the falling man to the death of the victim. I shall claim, shortly, that my account of justification and impartiality yields better results.

wrongly whatever one does. In their view, genuine moral dilemmas are irresolvable, since there is no *moral* consideration that can decide between the options with which one is confronted.⁵⁰

We are now in a position to see that, according to a plausible version of a two-tiered ethics of self-defense, the right to kill an innocent threat in self-defense might turn out to be a right to do wrong. Consider again the complaint against how Thomson explains the permission to kill an innocent threat in self-defense. 'A Nonresponsible Attacker is no more an agent than a tiger or a boulder, it seems that some [innocent attackers] and some [innocent threats] cannot violate rights....'⁵¹ As I have already pointed out, this argument is fallacious: persons, rather than tigers, might unintentionally violate their duties, hence the morality of interest generates a right to kill them in self-defense.

Still, there is a deep truth underlying this fallacious argument: by taking advantage of his right to self-defense, the potential victim treats the innocent threat *as if* he were a boulder rather than a person. Respecting him as a person involves considering the situation impartially; and from the perspective of the impartial observer, the interests of the innocent threat are as weighty as those of the victim. From this perspective, the fact that the innocent person poses a threat to the victim is irrelevant. Hence, even if the morality of interests allows it, the right to kill in self-defense might well be a right to do wrong; treating the falling man respectfully leads to a requirement to forego the moral right to kill him in self-defense. This is, however, not the end of the story. By the same token, from the impartial spectator's point of view, the victim's interest is as weighty as ever. Thus, when the victim waives his right to self-defense, he treats himself as a self-standing net, i.e., as a means for the survival of the innocently falling man. Hence, considered impartially, a case in which a person innocently threatens another person is a symmetrical case. As I have just noted, some philosophers believe that, in such a case, one acts wrongly whatever one chooses to do.

50 See David Wiggins, *Truth, Invention and the Meaning of Life* (Oxford: Oxford University Press, 1976), 376; W. Sinnott-Armstrong, "'Ought to Have' and 'Could Have,'" *Analysis* 45 (1985), 323-4; cf. R. Barcan Marcus, 'Moral Dilemmas and Consistencies,' *Journal of Philosophy* 77 (1980), 125. Thomson seems to disbelieve in moral dilemmas. See her 'The Right and the Good,' *Journal of Philosophy* 94 (1997), 284-6. Some philosophers, most notably Daniel Statman, deny the very possibility of such dilemmas. See his *Moral Dilemmas*. (Atlanta, GA: Rodopi 1995)

51 Jeff McMahan, 'Self-Defense and the Problem of the Innocent Attacker,' *Ethics* 104 (1994), 176

This has an obvious implication on the nature of the proportionality constraint by which the right to self-defense is governed. Killing the falling man in self-defense is unjustified, and in particular, it cannot be justified on the basis of the lesser evil principle. It follows that the proportionality by which the right of self-defense is constrained is not a rule of distributive justice. Thus, 'pointlessly' killing the culpable aggressor might be unjustified as well, because, in *terms of distributive justice*, it is disproportional. Still, the victim holds such a right; this is another case in which a right of self-defense turns out to be a right to do injustice. Usually, however, there is no such complication when killing a culpable aggressor is at stake. This is because it is fair that inevitable harms would be shifted to the villains, or to put it in Justice's terms, the villain is liable to harm. Impartiality is one of the essential virtues of just distributors; if the victim has a right to kill the culpable aggressor — if the Weberian Presumption was defeated — the impartial spectator would recommend taking advantage of this right.⁵²

This account of the asymmetry between innocent threats and villain aggressors neatly explains other data that cry out for theoretical illumination. First, as McMahan observes, commonsense morality unequivocally endorses a third-party intervention in other-defense just in cases where the attacker is culpable. I (partly⁵³) agree. The two-tiered ethics of self-defense embraces this intuition. In many cases the right to kill an innocent threat in self-defense is a right to do wrong. Hence, although the victim is permitted to kill in self-defense, he should give up this permission. Plausibly, Transferability is true only in cases in which killing in self-defense is justified. When the permission in question is a permission to act wrongly, it might not be transferable to others.

Second, consider the following case. Y is innocently threatening X's life. In defending himself, X is threatening Y's life. Thomson holds, strangely, that Y has no right to fight back. Thomson's commitment to this view follows from her belief (expressed in the Intuitive Argument)

52 It would be a mistake to conclude that killing an innocent threat in self-defense is *always* morally unjustified: if *one* person *innocently* threatens *ten* others, an impartial weighing would recommend preferring the ten; this would be harmful but not disrespectful to the one. The centrality of the realm of rights can be shown through a case in which the ten innocently threaten the one, and the one is able to kill the ten in self-defense. Killing the one in defense of the ten in this case is much more problematic.

53 I do not fully agree, because if one person innocently threatens ten people a third-party intervention is desirable. The reason for this is obvious in light of the discussion in the previous paragraph.

that X's having a right to act in self-defense implies that X is morally justified in doing so. If so, Y's attempt to stop X from committing a justified action cannot be justified. The two-tiered ethics of self-defense is free from this counterintuitive implication. If X's permission to kill in self-defense is a permission to act wrongly, it is perfectly possible that Y is permitted to defend himself against X's intentional aggression. Finally, consider another variant of an innocent-threat case. This time, the falling man who is about to kill you has just murdered someone. *Pace* Justice and W-Justice, it seems to me that, in this case, killing the threat in self-defense is morally justified. The reason provided by the two-tiered ethics of self-defense is clear enough. You have a right to kill the culpable threat in virtue of your interest to stay alive. And an impartial weighing of the interests at stake would lead you to use this permission. Again: it is fair that inevitable harms would be shifted to villains.

VII Conclusion

It has often been noted that the ethics of self-defense against innocent threats and the ethics of self-defense against villainous threats seem to be asymmetrical. This has been seen as a difficulty for Thomson's account of the ethics of self-defense, because that account recognizes no such asymmetry. My suggestion was to account for the asymmetry by reference to the distinction between a respect-based reason not to kill and a *material* interest-based reason not to kill. I propose that when one defends oneself against a villainous aggression, one is defending oneself against a threat of being killed disrespectfully, as well as against a threat of being materially harmed. But when one defends oneself against an innocent threat, one is only defending oneself against a threat of being materially harmed.

The argument for this claim has three major elements. First, I have shown that the right to kill an innocent person in self-defense is sensitive to how harmful the death of the victim would be — hence, the harm-based reason for self-defense. Second, I have shown the notion of liability used by Justice and W-Justice in order to explain the asymmetry between culpable and innocent aggressors fails to do that. Third, I suggested, in light of this result, that the role of culpability in justifying a right to self-defense should be explained by identifying a further interest the victim has when a culpable aggressor attacks him. When culpably attacked, one would suffer harmful death, which is also disrespectful. We saw that any plausible general morality of killing would be two-tiered in light of these two aspects of a death caused by a culpable aggression. Indeed, the two-tiered ethics of self-defense I have defended

is just an application of the two-tiered moral theory of the wrongness of killing.

Received: January 2004

Revised: September 2004

Revised: August 2005