

ExpressO Preprint Series

Year 2004

Paper 154

Fairness, Responsibility and Self-Defense

Re'em Segev
Harvard University

This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.

<http://law.bepress.com/expresso/eps/154>

Copyright ©2004 by the author.

Fairness, Responsibility and Self-Defense

Abstract

There is (almost) a consensus that self-defense is justified in some paradigm situations, but a substantial disagreement regarding the nature of this justification. The paper suggests a justification for self-defense which is based on a general thesis regarding the proper resolution of interpersonal conflicts. This thesis includes two basic ideas. First, individual well-being is the fundamental value. Second, interpersonal conflicts of well-being should be resolved in light of two conceptions of fairness, one of which considers the responsibility of persons for interpersonal conflicts as a morally significant factor. These ideas are elaborated in several principles that guide the resolution of interpersonal conflicts. The paper applies this thesis with respect to the justification of self-defense, explores its implications and compares it to other proposed justifications for self-defense.

Fairness, Responsibility and Self-Defense

Re'em Segev

Abstract

There is (almost) a consensus that self-defense is justified in some paradigm situations, but a substantial disagreement regarding the nature of this justification. The paper suggests a justification for self-defense which is based on a general thesis regarding the proper resolution of interpersonal conflicts. This thesis includes two basic ideas. First, individual well-being is the fundamental value. Second, interpersonal conflicts of well being should be resolved in light of two conceptions of fairness, one of which considers the responsibility of persons for interpersonal conflicts as a morally significant factor. These ideas are elaborated in several principles that guide the resolution of interpersonal conflicts. The paper applies this thesis with respect to the justification of self-defense, explores its implications and compares it to other proposed justifications for self-defense.

Table of Contents

I. Preface.....	2
II. Fairness as the Basis of Self-Defense.....	8
III. Different Fault-Based Justifications for Self-Defense.....	16
IV. Other Accounts of Self-defense.....	30
1. Rights.....	30
2. The Value of the Interests of Aggressors.....	39
3. Unjustified Invasions.....	44
V. Conclusion.....	50

 bepress Legal Repository

I. Preface

The content of normative rules should ideally be determined in light of their justification. Yet often we encounter a backward reasoning in which the content of norms is assumed to correspond to common intuitive judgments and the justification of these judgments is explored.¹ Nowhere perhaps is this backward reasoning so apparent than with respect to self-defense (including defense of property and defense of others).² There seems to be (almost) a consensus that self-defense is justified, at least in some paradigm situations (and this consensus is reflected in the law).³ Moreover, self-defense is typically considered the archetype justification for harming individuals. This view is accompanied by a significant agreement regarding the content and accordingly the boundaries of the justification for self-defense. Yet the justification for self-defense is typically assumed rather than accounted for.⁴ And, when it is explored, substantial disagreements emerge and many accounts seem unsatisfactory.⁵ The "repeated failure of moral and legal theories to establish... grounding" for self-defense has even led some to pessimism "concerning the whole project of attempting

¹ This kind of process is famously advocated by John Rawls in his proposal that normative conclusions should be determined through a method of "reflective equilibrium," which, among other things, explores the justification of "considered judgments." See John Rawls, *A Theory of Justice* (Revised Edition, Cambridge, Mass.: Harvard University Press, 1999), pp. 17-19, 42-45.

² I will use, from now on, for brevity, the term "self-defense," although my argument is applicable also to defense of property and defense of others. Unfortunately, common terminology concerning defensive force is problematic, especially in the second respect. The terms "self-defense" and "private defense" both imply that the justification for defensive force is agent-relative (varies in accordance with the identity of the agent): the former implies a distinction between an agent who defends his interests and an agent who defends another, and the latter implies a distinction between a private person and a public official. As I note in the next part, I believe that all relevant reasons and the conclusions they entail regarding the resolution of interpersonal conflicts are agent-neutral, namely, apply to all agents equally.

³ See the claim that "all Western legal system recognize" that self-defense is an exception to the prohibition on intentional killing: George P. Fletcher, "The Right to Life," *The Monist* 63, no. 2 (1980): 135-155, 139.

⁴ See Suzanne Uniacke, *Permissible Killing: The Self-Defense Justification of Homicide* (Cambridge: Cambridge University Press, 1994), p. 1.

⁵ Another context in a similar phenomenon seems to exist is punishment: there is a wide agreement that it is justified and a substantial disagreement regarding the nature of this justification.

to justify the act of defending (innocent) human life,"⁶ and others to consider self-defense as an excuse rather than a justification.⁷

I share the view that many accounts of the justification for self-defense are inadequate. It is also true that excuse is the proper account of many cases of self-defense: typically, a person can hardly be held responsible and, consequently, be blamed, for defending himself from an attack that might cause him (severe) harm.⁸ Yet, I nevertheless believe that some cases of self-defense could be rationally justified.

⁶ Michael Gorr, "Private Defense," *Law and Philosophy* 9, no. 3 (1990): 241-268, 264, 267-268 (suggesting that the justification for self-defense is basic and could not be accounted for).

⁷ Cathryn Jo Rosen, "The Excuse of Self-Defense: Correcting a Historical Accident on Behalf of Battered Women who Kill," *The American University Law Review* 36, no. 1 (1986): 11-56. It is unclear, however, whether her view is that self-defense is never justified or that it is sometimes justified but it is best to treat "all self-defense... as excused" for second order reasons relating to the adverse effects of self-help. The first order interpretation is supported by the claim that "there probably is no acceptable calculus to support treatment of self-defense as a justification" (p. 49). The second order interpretation is supported by the assertions that "few cases in which self-defense is claimed... fit the model of a justification" (p. 56); "treatment of battered woman's defense as an excuse does not preclude justifying women who kill men under objectively identifiable circumstances more akin to traditional self-defense" (p. 44); and "excused self-defense would better meet the needs of... society in general" (p. 45).

An ambiguous analysis of self-defense, with respect to its nature as a justification or excuse, is also found in William Blackstone, *Commentaries on the Laws of England* (1765) (Philadelphia: Childs & Peterson, 1860), book I, chapter 1, p. 130 ("Both the life and limbs of a man are of such high value, in the estimation of the law of England, that it pardons even homicide if committed *se defendendo*, or in order to preserve them. For whatever is done by a man to save either life or member, is looked upon as done upon the highest necessity and compulsion"), book III, chapter 1, pp. 3-4 (self-defense is lawful since "the law in this case respects the passions of the human mind, and... makes it lawful in him to do himself that immediate justice to which he is prompted by nature, and which no prudential motives are strong enough to restrain... Self-defence, therefore, as it is justly called the primary law of nature, so it is not, neither can it be in fact, taken away by the law of society. In the English law particularly it is held an excuse for breaches of the peace, nay, even for homicide itself"), book IV, chapter 14, p. 183 ("Homicide in *self-defence* or *se defendendo*, upon a sudden affray, is also excusable, rather than justifiable, by the English law.") This analysis is interpreted as reflecting an excuse rather than a justification by Hugo Bedau, "The Right to Life," *The Monist* 52, no. 4 (1968): 550-572, 559.

⁸ This point is emphasized, for example, by John Harris, *The Value of Life* (London: Routledge & Kegan Paul, 1985), pp. 67-68. McMahan argues that persons who kill innocent bystanders and innocent aggressors in order to save their lives cannot be excused. His reason is that some cases of this kind (such as that of terrorists) are not excused and others are no different in that in all there is the same threat to life. See Jeff McMahan, "Self-Defense and the Problem of the Innocent Attacker," *Ethics* 104, no. 2 (1994): 252-290, 287-288. But this argument is unfounded: the question whether a person should be excused is a function not of the nature of the threat but of the *effects* of the threat on him (the nature of the threat is only an indication for these effects) and these might be different in different cases of harming innocent bystanders. Typically, terrorists will be more responsible for their actions than persons who are caught in unexpected situations in which their life is in danger, and therefore the latter might typically be excused although the former will not.

The aim of this paper is to suggest such a justification for self-defense, in light of a general thesis I developed in several other papers regarding the proper resolution of interpersonal conflicts. This thesis includes principles that elaborate the ideas that individual well-being is the fundamental value and that internal clashes of independent aspects of it, namely, conflicts between the well-being of different individuals, should be resolved in light of two conceptions of fairness, one of which considers the responsibility of persons for interpersonal conflicts as a morally significant factor. This general framework provides, I believe, a plausible alternative for the common kinds of moral theories that dominate the inquiry of interpersonal conflicts: on the one hand, deontological theories and theories of rights and, on the other hand, aggregative forms of consequentialism, particularly utilitarianism.⁹ Interpersonal conflicts usually referred to under the heading of self-defense test a central aspect of this framework.

According to the view I suggest, the justification for self-defense is entailed by general principles that should guide the resolution of interpersonal conflict. Not all the principles I suggest apply in every interpersonal conflict, and the force of applicable principles varies in light of the nature of each conflict. But, on this view, self-defense is not a substantive category: no sharp line distinguishes between situations typically referred to under this heading from other interpersonal conflicts.

My inquiry explores the interplay of considerations of well-being and fairness, in the senses described in the next part, with respect to three issues relating to the resolution of interpersonal conflicts. The first, and most fundamental issue, is the nature of the justification for the paradigm case of self-defense, namely, a situation in which one person unjustifiably

⁹ This thesis is outlined in Re'em Segev, "Well-Being and Fairness: Beyond Deontology, Rights and Utilitarianism," and is further explored in Re'em Segev, "The Relevance of Numbers: Intrinsic and Comprehensive or Instrumental and Restricted?" In these papers, as in this one, I focus on small-scale, sharp and one-dimensional interpersonal conflicts involving basic interests, such as life, bodily integrity or property.

and culpably attacks and thereby endangers the well-being of another who is not responsible for the conflict.

The second issue I discuss is the implication of the basic justification for self-defense with respect to the justification for a requirement of proportionality. There is a consensus that it is justified to harm an aggressor in the paradigm case of self-defense, if this is necessary (as the only way) to stop the attack and thus prevent the danger it creates, even if the harm to the aggressor is more severe than the harm the attack creates.¹⁰ There is, however, a dispute as to whether the last proposition should be qualified by a requirement of proportionality, which excludes severely harming an aggressor in order to prevent the danger the attack creates to a trivial interest (for example, whether it is justified to kill an aggressor in order to prevent him from damaging a trivial property).¹¹ The thesis I suggest implies a requirement of proportionality.

Finally, I explore the relevance of the responsibility of individuals involved in interpersonal conflicts for the existence of the conflicts to their proper resolution. In this respect, I suggest that (1) culpable aggressors and others who are responsible for the existence of an interpersonal conflict should be distinguished from all other persons involved in interpersonal conflicts who are not responsible for their existence. This broad category

¹⁰ A standard example is of a person who defends himself by harming *several* aggressors. I ignore this example since it is based on the dubious assumption that harm to several persons is worse than an identical harm to one person. I question this assumption in "The Relevance of Numbers." I thus focus in conflicts involving only one person on each side.

¹¹ The requirement that the force used will be necessary in order to stop the attack and the requirement of proportionality are independent. Suzanne Uniacke argues that "Although necessity and proportionality are distinguishable conditions... considerations of proportionality... must sometimes form part of the normative background against which necessity is judged." But the examples she gives are of cases in which force is strictly necessary, without regard to consideration of proportion. In the first kind of examples – in which "instead of using lethal force against an aggressor I could avoid being killed myself by complying with the aggressor's demands, (say) by participating in a murder or by revealing the whereabouts of someone who will be then endangered" – although the force considered is not necessary in order to save an interest of the agent, it is strictly necessary in order to save the interest of another person. The second kind of examples – in which the endangered agent could avoid being in the dangerous situations – does not demonstrate that force is not strictly necessary, but rather that other considerations – concerning prior responsibility – might also be relevant. See Uniacke, *Permissible Killing*, pp. 82-83.

includes: (2) "Innocent aggressors" - who act with the intention to harm others but are not responsible for their behavior, for example, an insane aggressor; (3) "Innocent threats" - who constitute physical threats without acting, for example, a person who is thrown by another down a deep well at the bottom of which there is a third person who will be crushed and killed by the body of the falling person; (4) "Innocent shields of threats" - who are "so situated that they will be damaged by the only means available for stopping the threat," for example, a person who is strapped to the front of a tank of an aggressor so that the tank cannot be hit without also harming him¹²; (5) "Bystanders" – who are not related to the danger in any way but sacrificing or harming them is the only way to save others, for example, each of two drowning persons only one of which could be saved, or a person that harming his interest is the only way to save the interest of another person. (I thus use the term "innocent" in the only sense I consider morally significant, namely, not responsible for the conflict (not culpable), rather than in the technical sense, often used, of not (wrongfully) constituting a threat.¹³)¹⁴

The analysis of the justification for self-defense and its boundaries is essentially applicable both morally and legally. While the law should not necessarily reflect every moral conclusion, this is mainly since some moral conclusions could not be significantly furthered by their incorporation into the law, or are not important enough to justify their incorporation into the law, in light of the price of legal enforcement. These considerations do not apply, or are not

¹² The terms and quotes concerning the last two categories are from Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), pp. 34-35 (who suggests that a self-defense justification is applicable to innocent threats and leaves open the question of innocent shields of threats).

¹³ For the technical sense see, for example, Judith Jarvis Thomson, "Rights and Deaths," *Philosophy & Public Affairs* 2, no. 2 (1972): 146-159, 154; Uniacke, *Permissible Killing*, pp. 78-81, 95-96.

¹⁴ I do not discuss other aspects of the justification and content of self-defense since they are not directly relevant to the general principles I suggest. For example, whether, in light of the requirement of necessity, (safe) retreat is generally required or whether a self-defense justification requires that the agent's actions be accompanied with a belief (or knowledge) that the necessary factual conditions are present and perhaps an intent to defend himself or someone else. With respect to the last question, I assume that at least belief is necessary. This issue is discussed, for example, by Russell Christopher, "Self-Defense and Defense of Others," *Philosophy and Public Affairs* 27, no. 2 (1998): 123-141.

decisive, in the context of interpersonal conflicts involving basic interests. (There might be special importance to clarity, even at the expense of precision, in formulating norms addressed to many persons, such as legal norms, but this - moral - consideration is relevant also to moral norms when they could influence many people.) Indeed, although there is a disagreement as to the proper limits of law, it is widely agreed that it should at least reflect moral judgments with respect to the protection of important interests of people from harm.¹⁵ This idea presents various difficult questions regarding its proper content and boundaries,¹⁶ but its hard core seems obviously justified and widely accepted, while the main dispute is whether the law should go beyond it.¹⁷ Thus, the analysis of the proper resolution of interpersonal conflicts in general, and the moral justification of self-defense in particular, is directly applicable in the legal sphere.¹⁸

¹⁵ The classical articulation of the position that the prevention of harm - to people other than the agent - is the only justified aim for the use of coercion, including by law, is John Stuart Mill, *On Liberty* (1859) (David Bromwich and George Kateb, eds., New Haven: Yale University Press, 2003).

¹⁶ For a brief survey of some of the main problems of Mill's Harm Principle see John Gray, "Introduction" to John Stuart Mill, *On Liberty and Other Essays* (Edited with an Introduction and notes by John Gray, Oxford: Oxford University Press, 1998), pp. xv-xxi.

¹⁷ The justification of other grounds of invoking the law is discussed by H. L. A. Hart, *Law, Liberty and Morality* (Oxford: Oxford University Press, 1963); and Joel Feinberg, *The Moral Limits of the Criminal Law: vol. 1 - Harm to Others* (New York: Oxford University Press, 1984); *vol. 2 - Offense to Others* (New York: Oxford University Press, 1985); *vol. 3 - Harm to Self* (New York: Oxford University Press, 1986); *vol. 4 - Harmless Wrongdoing* (New York: Oxford University Press, 1988).

¹⁸ Kent Greenawalt, while emphasizing the connection between moral and legal standards (he refers mainly to accepted legal norms but clearly considers proper moral and legal standards to be closely connected), inter alia, in the context of self-defense, nevertheless suggests that there are distinctions between them beyond those I mentioned above. For example, he suggests that even if there is a moral distinction between culpable and innocent aggressors, it should not necessarily be reflected in the law, since most aggressors are culpable and "many cases of innocent aggression will reasonably appear wrongful to the victim, who must, of course, act quickly without much thought about the character of his assailant." Based on this observation, he writes, "One can well understand why legal standards are not drafted to require assessment of the victim's knowledge of an aggressor's innocence in these exceptionally rare cases." Kent Greenawalt, "Violence - Legal Justification and Moral Appraisal," *Emory Law Journal* 32, no. 2 (1983): 437-497, 474. Greenawalt's factual observation may well be correct, but I fail to see how it justifies a difference between moral and legal standards in this context. This observation is relevant to the evaluation of the justification of actions both in the legal and in the moral sphere. If there is a distinction between culpable and innocent aggressors and the assumptions of agents in this respect are relevant to the justification of their actions (as Greenawalt assumes), why not reflect this distinction in the law, even if it will be relevant only in rare cases?

I will start, in the next part, by with an outline of the general principles I propose for the resolution of interpersonal conflicts and then explain in general and then explain how they apply to self-defense. In part III, I compare my analysis, and especially the responsibility-based principle I suggest, to several other principles that resemble it but are nevertheless different in other respects and explain how I confront various objections made against these principles. Finally, in part IV, I believe my suggestion is preferable to different kinds of dominant justification for self-defense.

II. Fairness as the Basis of Self-Defense

The general thesis I suggest for the resolution of interpersonal conflicts reflects two basic ideas.¹⁹ First, the fundamental value and source for reasons for actions, in the context of interpersonal conflicts, is individual well-being. Second, the resolution of interpersonal conflicts - in which there is an internal clash of reasons for action to protect the well-being of different persons – should be done in light of conceptions of fairness that reflect the intrinsic and independent value of each person and the rational power of persons. These ideas are elaborated in six principles, all reflecting agent-neutral (impartial) reasons for action:

(1) *The Well-Being Principle* reflects the intrinsic value of each person: there is a reason to protect, and perhaps enhance, the well-being of persons.

(2) *The Importance Principle* reflects the difference in the importance of various aspects of individual well-being: the strength of the reason provided by *The Well-Being Principle* depends on the importance of the interest at stake, the probability that it will be affected (for better or for worse) if an action is not performed and the probability that an action will affect it. (*The Importance Principle* thus elaborates and absorbs *The Well-Being Principle*.)

¹⁹ These ideas, and the principles they entail, are elaborated in "Well-Being and Fairness."

Assuming that all relevant probabilities are equal, the more important is the interest, the stronger is the reason to protect it. In resolving interpersonal conflicts, there is, therefore, a reason to prefer the person who would be affected most significantly by the resolution of the conflict (for example, the person whose life would be saved over the person whose property would be saved).

(3) *The Equal Chance Principle* addresses clashes of reasons generated by *The Well-Being Principle*. It reflects the independent value of each person and the ensuing intrinsic importance of each person's fate; the fact that a resolution of an interpersonal conflict is for each of the persons whose interests are involved in the conflict a question of all or nothing (as far as the interest at stake is concerned); and the lack of common denominator in light of which interests of different persons could be traded-off without a loss. In light of these truths, this principle, which I consider an aspect of the concept of fairness, holds that there is a reason to accord equal weight to the well-being of each person who will be affected by the resolution of an interpersonal conflict. This is done (roughly) by choosing which person to prefer in a way that gives each person an equal chance.

(4) *The Qualitative Difference Principle* resolves the clash of the last two principles. The reason provided by *The Importance Principle* prevails over the reason provided by *The Equal Chance Principle* if, and only if, assuming that all relevant probabilities are equal, there is a qualitative gap in importance of the competing interests.

(5) *The Principle of Fairness: Responsibility*, a second aspect of the concept of fairness, reflects the rational power of persons by making them accountable for their choices. It holds that the responsibility of a person for an interpersonal conflict is a reason that he, rather than another person, would bear the burden or enjoy the benefit involved in the resolution of the conflict. (Interpersonal conflicts might involve burdens, benefits or both.) The resolution of clashes between this reason and the reasons provided by the previous principles depends on

the relative force of the clashing reasons, namely, on the importance of the competing interests, the relevant probabilities and the degree of responsibility of each person for the existence of the conflict.

These principles are, in several respects, both different from common conceptions and counterintuitive. I will not repeat here my attempt to defend them in their general form. Rather, my aim here is to reinforce this defense by demonstrating that they offer a plausible justification for self-defense. Several aspects of these principles are not directly relevant for the exploration of the justification for self-defense. I mention all principles in order to clarify the assumptions underlying the following discussion, but I will mostly ignore the aspects that are not directly relevant for my argument.

These principles entail the following conclusions with respect to the resolution of interpersonal conflicts in general and conflicts typically referred to as self-defense situations, especially the paradigm case of self-defense, in particular. In interpersonal conflicts, in which a choice has to be made between the well-being of two or more persons, there is a reason, or rather contrary reasons, expressed in *The Well-Being Principle*, to take account and protect the well-being of each person. The force of each of these reasons is determined in light of The Importance Principle. When no one is responsible for the conflict, we should follow The Equal Chance Principle, namely, accord each person an equal chance to be preferred, unless there is qualitative difference in the importance of the competing interests, in which case The Qualitative Difference Principle requires preferring the person with the qualitatively more important interest.²⁰

The analysis is more complex when one or more persons are responsible for the conflict. The reason, reflected in *The Well-Being Principle*, to take account of the interests of each person, applies also with respect to the interests of persons who are responsible for

²⁰ I elaborate on this issue in "Well-Being and Fairness."

interpersonal conflicts, including unjustified and culpable aggressors in the paradigm case of self-defense. An aggressor, even an unjustified and culpable aggressor, is still a person that his well-being should be taken into account. Therefore, the above analysis applies also in the paradigm case of self-defense, subject to one important qualification: the responsibility of a person for an interpersonal conflict, namely, for the need to make a choice between the interests of several persons, provides another - contrary - reason, stated in *The Principle of Fairness: Responsibility*, to prefer the non-responsible person, namely, to harm the responsible person in order to save the non-responsible person. (I assume throughout that the person who is being attacked is in no way responsible for his situation. If he is, there is another reason against preferring him. The force of the reasons based on *The Principle of Fairness: Responsibility* varies in accordance to the degree of responsibility.)

The resolution of interpersonal conflicts in which one person is responsible for the existence of the conflict, and particularly of the paradigm case of self-defense, should be thus as follows. When the competing interests are of equal importance, *The Well-Being Principle* provides two contrary reasons of equal importance, and *The Equal Chance Principle* thus provides a reason to accord each an equal chance to be preferred. However, this reason is outweighed by the reason reflected in *The Principle of Fairness: Responsibility*, to prefer the non-responsible person, if the degree of responsibility is significant (not trivial). Since this condition is met (by hypothesis) in the paradigm case of self-defense, *The Principle of Fairness: Responsibility* is thus the basis for the justification to harm the aggressor who is responsible for the conflict in order to save the non-responsible person.

The situation might be more complicated when the interests of the aggressor and the non-responsible person are not equally important, since then the reason provided by *The Importance Principle* should also be taken into account. The conclusion is obvious when the choice is between the less important interest of the aggressor and the more important interest

of the non-responsible person, since then the reason provided by *The Importance Principle* reinforce the reason provide by *The Principle of Fairness: Responsibility* and both advocate the preference of the same (non-responsible) person.

The situation is more complicated when the aggressor's interest is more important than that of the non-responsible person, since then the reason, provided by *The Principle of Fairness: Responsibility*, to prefer the non-responsible person, clashes with the reason, provided by *The Importance Principle*, to prefer the more important interest of the aggressor. For example, when an aggressor threatens the property or bodily integrity of another and the only way to prevent this is by killing the aggressor.

Each of the competing principles might prevail in this kind of conflict, depending on its force relative to the force of the competing principle in each conflict. When the reason provided by *The Principle of Fairness: Responsibility* is very strong, as in the paradigm case of self-defense, this reason will prevail over the reason provided by *The Importance Principle*, as long as the latter reason is not very strong as well. In other words, *The Principle of Fairness: Responsibility* will prevail over *The Importance Principle* even if the aggressor's interest is more important than that of the non-responsible person provided that the difference in importance does not reach a certain threshold.

What is this threshold? With respect to conflicts between the interests of two non-responsible persons, I suggested *The Qualitative Difference Principle: The Importance Principle* should prevail over *The Equal Chance Principle* when the difference in the importance of the competing interests is qualitative. In situations in which one person is responsible for the conflict, so that *The Principle of Fairness: Responsibility* provides an additional reason to prefer the non-responsible individual, the difference in important must be bigger if *The Importance Principle* is to prevail. The resolution of the clash between *The Principle of Fairness: Responsibility* and *The Importance Principle* thus entails a principle of

proportionality in self-defense situations. According to this principle, it is justified to harm an aggressor in order to save a non-responsible person even if the only way to do this requires causing the aggressor more harm than the harm he will otherwise cause to another, as long as the difference is not very big.

It is hard to determine what should be the exact boundaries of the proportionality principle, but it seems that it entails, *inter alia*, the following conclusions with respect to the paradigm case of self-defense. It is unjustified to kill or severely injure an aggressor is unjustified in order to prevent him from trivially, or even significantly, damaging the property of another, and even from causing a trivial bodily injury to another, such as a scratch. It is, however, justified to injure severely, and even to kill, an aggressor who would otherwise cause a substantial bodily injury to another. It is also justified to destroy a valuable property of an aggressor if he would otherwise destroy a relatively minor property of another. But it might be unjustified to destroy an aggressor's house in order to prevent him from causing trivial damage to another's property.

The Principle of Fairness: Responsibility is thus the heart of the justification to harm an aggressor in the paradigm case of self-defense. I will therefore end this part with by noting several points concerning its basis and scope.

Note, first, that *The Principle of Fairness: Responsibility* does not reflect a retributive idea, according to which there is a reason to accord a reward for a right action and a penalty for a wrong action. Rather, *The Principle of Fairness: Responsibility* holds that when a choice has to be made between the interests of several individuals, there is a reason that a person who is responsible for the conflict should bear this cost rather than a non-responsible person.

Accordingly, *The Principle of Fairness: Responsibility* applies with respect to every person who is responsible for the existence of an interpersonal conflict, and this includes not just persons that their responsibility consists of performing wrongful acts but also persons whose

responsibility constitutes in taking of special risks that are justified.²¹ For example, suppose that a person engages in a risky recreational activity, such as swimming, and is caught in a storm so that the only way to save his life is by using a boat in a way that will damage the boat substantially. If other factors are equal, *The Principle of Fairness: Responsibility* holds that it is right to use the swimmer's boat, rather than a boat of another, even if we assume that the risk he took was justified in light of the dangers involved compared to the expected advantages. Similarly, *The Principle of Fairness: Responsibility* explains why an individual is not required to save another from a risk of harm that is the outcome of the other responsibility at the expense of exposing himself to an identical risk. Suppose, for example, that a person engages in a recreational activity which involves some risks and consequently is in danger of being harmed, and the only way to save him involved an identical risk for the saver. If the person at risk were not responsible for the situation, then *The Equal Chance Principle* would apply, since *The Well-Being Principle* would generate identical but contrary reasons to save one and not risk the second. But since the endangered person is responsible for the situation, the reason to prefer the other person becomes decisive and he is thus not required to risk himself in order to save the person in danger.

The Principle of Fairness: Responsibility refers to the responsibility for the threat and not its proximity, in terms of either time or place. For this reason, *The Principle of Fairness: Responsibility* applies with respect to persons who are responsible for the existence of interpersonal conflicts even if they are not the direct or immediate source of the threat or are not in the vicinity of the threat. For example, this principle applies to persons who coerce others to harm third parties and to persons who are responsible for a situation in which a choice needs to be made between their interests and those of others. (Although it might often be impossible to resolve the conflict by harming such persons, this might be occasionally

²¹ I explain the reasons for this position in "Well-Being and Fairness."

possible.) In other words, *The Principle of Fairness: Responsibility* is applicable to not only responsible aggressors but also to what might be called responsible "bystanders," namely, responsible person who do not constitute a current or direct threat. (Whether or not we refer to the justification for harming such persons, based on *The Principle of Fairness: Responsibility*, under the category of self-defense is, of course, not important.)

The Principle of Fairness: Responsibility does not, however, apply with respect to persons who are not responsible for the existence of an interpersonal conflict even if they constitute threats or are part of threats (including "innocent aggressors," "innocent threats" or "innocent shields of threats"). While there is a reason to prefer, when resolving an interpersonal conflict, the non-responsible person over the one responsible for the conflict, there is no reason to prefer one person over another just because one constitutes a threat or is part of a threat, assuming that this he is not responsible for this. In other words, there is no reason to distinguish between non-responsible (innocent) persons, whatever their role in the conflict: bystanders, aggressors, threats or shields of threats. In fact, since there is no substantive difference between various categories of innocent persons, the lines between these categories are hard to draw on non-arbitrary grounds. (It is important to remember, however, that the conclusion that it is unjustified to harm one innocent person in order to save another, even if the former is an innocent aggressor, threat or shield of threat, does not necessarily entail the conclusion that the defender is responsible or blameworthy for doing so. He might be excused.)

Finally, there might be situations in which an act of defense against an unjustified and responsible aggressor will have collateral effects of harming (or endangering) non-responsible (innocent) persons (whether they are innocent bystanders, innocent threats or innocent shields of threats). The above analysis leads to the conclusion that a defensive act is justified in such situations only if the interests of the (non-responsible) person under attack are qualitatively

more important than the interests of each of the other non-responsible (innocent) persons who will be affected by this act (so that *The Qualitative Difference Principle* outweighs *The Equal Chance Principle*).

III. Different Fault-Based Justifications for Self-Defense

In this part, I discuss several proposed justifications for self-defense that are similar to mine in that they consider the responsibility for the conflict as a crucial factor. I will explain what are the differences between my thesis and these views and explore several objections made to these views that might be relevant to my thesis.

The idea that culpability or responsibility is a relevant factor in the justification of self-defense was raised by several writers (who consequently equate innocent aggressors and innocent threats, as well as innocent shields of threats, with innocent bystanders).²² Phillip Montague, George Draper and Jeff McMahan suggest accounts of self-defense which are especially close to *The Principle of Fairness: Responsibility*, although they differ from it in several important respects. Montague's idea is that in situations in which some of several individuals will be unavoidably harmed, it is right to sacrifice the one "whose fault it is" that this situation exists.²³ In a later article Montague described this idea as follows: "when unavoidable harm in being distributed among a group of individuals, and when some members of the group are to blame for the predicament of all, then justice requires (ceteris

²² See, in addition to the ideas I explore in detail below, Larry Alexander, "Self-Defense, Justification, and Excuse," *Philosophy and Public Affairs* 22, no. 1 (1993): 53-66; Michael Otsuka, "Killing the Innocent in Self-Defense," *Philosophy and Public Affairs* 23, no. 1 (1994): 74-94. For the moral equivalence of various categories of non-responsible persons see also Nancy Davis, "Abortion and Self-Defense," *Philosophy and Public Affairs* 13, no. 3 (1984): 175, 188-194; McMahan, "Self-Defense and the Problem of the Innocent Attacker," pp. 266-268.

²³ Phillip Montague, "Self-Defense and Choosing between Lives," *Philosophical Studies* 40, no. 2 (1981): 207, 215.

paribus) that the harm be distributed among those who are blameworthy."²⁴ Similarly, Draper suggests the following principles. First, "other thing being equal, if only one individual is (morally) to blame for the fact that someone must sustain a cost, then that individual should sustain the cost." Second, "Other thing being equal, if more than one person is to blame for the fact that someone must sustain a cost, then that cost should be divided among those persons in proportion to their blame or, if the cost cannot be divided, then the one who is most to blame should sustain the cost."²⁵ Finally, McMahan suggests that "in cases in which a person's culpable action... has made it inevitable that someone must suffer harm, it is normally permissible, as a matter of justice, to ensure that it is the culpable person who is harmed rather than allowing the cost of his wrongful action to be imposed on the morally innocent."^{26 27}

All these principles are subject to similar qualifications. The first is obvious as it ensures that an interpersonal conflict indeed exists: the harm must be necessary.²⁸ The second qualification is less obvious, and its basis is hardly explained by these writers, but, as I

²⁴ Phillip Montague, "Punishment and Societal Defense," *Criminal Justice Ethics* 2, no. 1 (1983): 30-36, 32. See also Phillip Montague, *Punishment as Societal Defense* (Lanham, Maryland: Rowman & Littlefield, 1995), pp. 34-50.

²⁵ George Draper, "Fairness and Self-Defense," *Social Theory and Practice* 19, no. 1 (1993): 72-92, 76-77.

²⁶ McMahan, "Self-Defense and the Problem of the Innocent Attacker," p. 259.

²⁷ These accounts seem to lead to the conclusion I suggested in the previous part, namely, that it is preferable to resolve interpersonal conflict by harming the persons responsible for their existence even if they are not the direct source of danger (such as those who coerce others to harm third parties and are responsible for a situation in which a choice needs to be made between their interests and those of others). See McMahan, "Self-Defense and the Problem of the Innocent Attacker," p. 258; Draper, "Fairness and Self-Defense," pp. 80-81; Kai Draper, "Self-Defense, Collective Obligation, and Noncombatant Liability," *Social Theory and Practice* 24, no. 1 (1998): 57-81, 63-68 (interpreting responsibility to include failure to prevent a conflict but only if the agent has an "obligation" to do so). For a similar view on this point see Lawrence Alexander, "Self-Defense and the Killing of Noncombatants: A Reply to Fullinwider," *Philosophy and Public Affairs* 5, no. 4 (1976): 408, 410-412; Paul H. Robinson, *Criminal Law Defenses* (Minnesota: West, 1984), volume 2, p. 73.

²⁸ Montague, "Punishment and Societal Defense," p. 33; Montague, *Punishment as Societal Defense*, p. 46; Draper, "Fairness and Self-Defense," pp. 78-79; McMahan, "Self-Defense and the Problem of the Innocent Attacker," p. 262.

explained in the previous part, I believe that it is justified: a requirement of proportionality.²⁹

The last qualification limits these principles so that they justify harming only culpable individuals.³⁰

While these ideas are close to my suggestion of *The Principle of Fairness: Responsibility*, there are nevertheless several important differences. First, I believe that a proper resolution of all interpersonal conflicts must rest on an analysis of several general (possibly) competing reasons for action, as reflected in the six principles I outlined in the previous part. The above analyses are narrower. They focus almost exclusively on the fault-based principles and although they mention other - including competing - considerations, and qualify these principles (in the ways just mentioned) in order to accommodate them, they do so only in few words that hardly explain what is the source of these clashes and how to resolve clashes between their principles and other considerations. (They avoid this by focusing mainly on one kind of interpersonal conflict: conflicts between equally important interests - mainly lives - of several individuals that some of which are responsible for the situation.)

²⁹ Montague, "Punishment and Societal Defense," p. 33; Montague, *Punishment as Societal Defense*, pp. 45-46 (adopting "a proportionality condition according to which the distribution of unavoidable harm among those who are to blame for the existence of that harm must be proportional to the harm that would be suffered by innocent person under a different distribution"; This condition is interpreted to exclude doing "major harm" to a blameworthy individual in order to prevent a "minor harm" to an innocent individual); Draper, "Fairness and Self-Defense," pp. 79-80 ("a consideration of welfare overrides the consideration of fairness in this case... proportionality requires only that the harm inflicted be comparable to (or less than) the harm thereby avoided"); McMahan, "Self-Defense and the Problem of the Innocent Attacker," footnote 11 at p. 259 ("it may not be permissible to inflict a great harm on a person in order to prevent his culpable action from causing a trivial harm").

³⁰ Montague, "Punishment and Societal Defense," pp. 32-3; Montague, *Punishment as Societal Defense*, p. 46 (stating that his principle does not "straightforwardly" apply to situations in which harming a blameworthy individual involves also harming an innocent individual); Draper, "Fairness and Self-Defense," pp. 80-81; Draper, "Self-Defense, Collective Obligation, and Noncombatant Liability," pp. 59-63 (suggesting a "divide-and-conquer strategy" that justifies a single action based on several principles: some of the harmful effects of the action are justified by self-defense while other harmful effects are justified by other principles, such as the principle of double effect or the principle that it is justified to harm innocent persons if this will prevent much more severe harm to other innocent persons; the main formulation of the last principle refers to the number of innocent persons harmed, but this principle is valid regardless of the question whether the numbers count).

Moreover, the position of Montague and McMahan with respect to the one clash they both mention, between their fault-based principles and other principles, is unclear. On the one hand, both state that their principles should be subject to a requirement of proportionality, so that an important interest of a responsible party may be preferred to a minor interest of a non-responsible party. On the other hand, both seem to understand their principles as absolute in this respect in various contexts. McMahan writes that if his principle is accepted within a consequential theory, it should have "absolute priority in the evaluation of outcomes" by outweighing "all other features of the possible outcomes." Thus, he claims, a defender is justified in killing a culpable aggressor even if the probability that the threat he poses would materialize "is relatively low" and in killing "any number" of culpable aggressors.³¹ Similarly, Montague seems to consider it justified to harm the persons at fault for an interpersonal conflict "no matter how many" such persons are harmed as a result.³² (I am not sure that the number of individuals involved in interpersonal conflicts is in itself morally significant,³³ but this is the assumption of McMahan and Montague.³⁴) These assertions seem incompatible with a principle of proportionality. My view, as I explained in the previous part, is that *The Principle of Fairness: Responsibility* states a reason for action, which might clash with other reasons, and particularly with *The Importance Principle*, and that this is the basis for a requirement of proportionality.³⁵

Another difference concerns the foundation of the idea that, other things being equal, it is right that the person responsible for an interpersonal conflict should bear its cost. Mostly, this

³¹ McMahan, "Self-Defense and the Problem of the Innocent Attacker," p. 261.

³² Montague, "Self-Defense and Choosing between Lives," p. 215; Montague, *Punishment as Societal Defense*, p. 43.

³³ I discuss this issue in "The Relevance of Numbers."

³⁴ McMahan makes this claim explicitly in this context: McMahan, "Self-Defense and the Problem of the Innocent Attacker," pp. 260-261. Montague explicitly consider the numbers as relevant in interpersonal conflicts that are no one's fault: Montague, *Punishment as Societal Defense*, p. 39.

³⁵ I explained how I think clashes between *The Principle of Fairness: Responsibility* and other principles should be resolved more generally in "Well-Being and Fairness."

issue is not discussed in the above accounts. The only brief reference is McMahan's remark that his variation of this idea "is best interpreted" and in fact "must... be understood as a deontological theory." (He rejects several consequential interpretations of his principle.) He points to two matters in which his principle is deontological. First, he considers the sacrifice of the culpable party as a just action and not as an action that is right "because it produces a better or more just outcome." Second, he determines who is a culpable (unjustified and unexcused) party in light of deontological ideas such as the prohibition on intentionally killing an innocent person in order to prevent the intentional killing of several innocent persons.³⁶ My thesis, including *The Principle of Fairness: Responsibility*, rests on a different foundation, which is suggested as an alternative to deontological and rights-based theories, on the one hand, and to common forms of consequentialism, on the other. As I explained, the principles I suggest elaborate two basic ideas: that the fundamental value and source for reasons for actions, in the context of interpersonal conflicts, is individual well-being; and that the resolution of interpersonal conflicts should reflect the independent value of each person and the rational power of persons. None of the principles I propose, including those that reflect the second idea, and particularly *The Principle of Fairness: Responsibility*, deviates from the first idea that the fundamental value is individual well-being. Rather, these principles, including *The Principle of Fairness: Responsibility*, complete this idea, by pointing to the proper – fair – way of resolving internal clashes of independent aspects of it, which occur when interpersonal conflicts of well-being emerge. Thus, contrary to McMahan, the principles I suggest do not reflect a deontological view.³⁷

³⁶ McMahan, "Self-Defense and the Problem of the Innocent Attacker," pp. 260-263. Similarly, Otsuka adheres to deontological distinctions in this context, such as between "killing and letting die," "initiation" and "redirection" of dangers and "using" a person and other means of harming a person: Otsuka, "Killing the Innocent in Self-Defense," pp. 76-79, 86-88.

³⁷ I discuss the differences between my thesis and dominant moral theories, mainly standard deontological and rights-based theories, on the one hand, and common forms of consequentialism, on the other, in "Well-Being and Fairness."

This difference leads to another difference regarding the background for the relevant principles. McMahan suggests that the background principle, in the context of interpersonal conflicts, should be a "presumption against shifting harms," which is based on his (deontological) assumption that there is a moral difference between doing harm and allowing harm.³⁸ I believe, on the other hand, that this distinction is unfounded,³⁹ and consequently that the background principle in the context of interpersonal conflicts should be *The Equal Chance Principle*, which rejects a presumption against shifting harms.

The reason, stated by *The Equal Chance Principle*, to accord each person involves in an interpersonal conflict is often overlooked. Many seem to consider a choice between the equally important interests of non-responsible persons as a matter of indifference. It is therefore important to emphasize that only this reason could be outweighed only by strong reasons. For example, when two people need an object in order to survive, the mere fact the one of them is the owner of the object may not provide a basis for a reason strong enough to outweigh *The Equal Chance Principle*, as suggested by Draper.⁴⁰ (Note, however, that ownership might often be related to other relevant considerations, for example, if the owner of the object invests resources in acquiring it exactly for this purpose, in which case *The Principle of Fairness: Responsibility* provides a reason to prefer giving the object to him.)

Another important difference between my *Principle of Fairness: Responsibility* and the principles suggested by Montague and McMahan concerns their scope. *The Principle of Fairness: Responsibility* refers to responsibility for an interpersonal conflict and thus might apply, as I explained in the previous part, not only to wrongful actions but also to justified

³⁸ McMahan, "Self-Defense and the Problem of the Innocent Attacker," pp. 252-253. Similarly, Otsuka's view is that if a danger to one person can be prevented only by harming another it is unjustified to act: Otsuka, "Killing the Innocent in Self-Defense," pp. 76-77 (especially footnotes 7 and 9).

³⁹ "Well-Being and Fairness."

⁴⁰ Draper, "Fairness and Self-Defense," pp. 84-85; Kai Devlin, "Justice as a Basis for Tort Liability," *Public Affairs Quarterly* 8(3) (1994): 237-253, 241-242.

actions.⁴¹ On the other hand, both Montague and McMahan limit their principles to wrongful actions. Montague refers to "fault" or to "culpable" actions and explicates these terms as including "intentional, reckless or negligent" actions, namely, unjustified (and perhaps also unexcused) actions, that generate interpersonal conflicts.⁴² Similarly, McMahan's principle is explicitly limited to morally "culpable" parties, which again seem to refer to unjustified (and unexcused) actions.⁴³ (McMahan explores several explanations for the intuitive assumption that there is a different, supplementary, justification, more limited in nature, for preferring innocent victims to innocent aggressors and innocent threats, but not to innocent bystanders. But he is "unable to find a fully convincing justification" for this assumption. He suggests that it might be due to the tendency of "overgeneralizing from the paradigm" of a culpable aggressor and concludes that this "reveals both an incoherence in commonsense morality and a lack of grounding for an important set of discriminations in the law," which, he assume, reflects this assumption.⁴⁴)

Draper's position in this respect seems closer to mine. Although he initially refers to "moral fault" or "blame," he later expands his account by suggesting the following principle: if "someone justifiably and for his own benefit creates a situation in which he or another innocent person must sustain a cost... fairness prefers that the cost should fall on the individual who created this situation." However, Draper seems to intend this formula to apply more broadly than I think justified by considering an effort to avoid a danger one is not responsible for as a benefit. Thus, his example is striking an undercover police officer who is

⁴¹ However, the force of this principle is affected by the degree of responsibility for the conflict which, in turn, might be affected by the question whether, or to what extent, the relevant action is justified.

⁴² Montague, "Self-Defense and Choosing between Lives," p. 211, 219; Phillip Montague, "The Morality of Self-Defense: A Reply to Wasserman," *Philosophy and Public Affairs* 18, no. 1 (1989): 81-89, 84; Montague, *Punishment as Societal Defense*, pp. 40, 44-45.

⁴³ McMahan, "Self-Defense and the Problem of the Innocent Attacker," pp. 259, 263.

⁴⁴ McMahan, "Self-Defense and the Problem of the Innocent Attacker," pp. 263-290.

trying to strike the defender since this is the only way to save his life.⁴⁵ Unless the police officer is responsible for the initial danger, I do not think that there is a reason to prefer either.

(Another difference regarding the scope of the above principles is that I consider *The Principle of Fairness: Responsibility* as applicable both to benefits and to burdens, while the other principles refer only to the allocation of burdens.)

Finally, other differences seem to be more general in nature. I believe that the principles suggested above should guide all agents in the resolution of all kinds of interpersonal conflicts, while Montague and McMahan make several distinctions in these respects. Montague claims that in standard self-defense situations one is "not only permitted to kill in self-defense, but one has a right to do so" and a third party is "obliged to choose in favor of the person whose life is threatened." And, he argues further, in other interpersonal conflicts "one is merely at liberty to choose in favor of himself" and a third party is "free to choose as he wishes under certain conditions."⁴⁶ McMahan similarly claims that self-defense should be "optional," namely, that "one is not required" to protect himself against a culpable aggressor.⁴⁷ These distinctions reflect complex issues that are beyond the scope of this paper. I will therefore just note that I believe that none of these distinctions reflects a substantive difference: the same principles should govern all kinds of interpersonal conflicts and apply to

⁴⁵ Draper, "Fairness and Self-Defense," pp. 83-84. See also his claim that individuals who create risks to others should compensate them if the risks materialize even if the creation of the risk was justified (not blameworthy): Devlin, "Justice as a Basis for Tort Liability," pp. 246-247. Compare to Otsuka's claim that it is justified to harm a person who is responsible for a threat on another but is innocent of blame, for example, since he mistakenly thinks that the other is about to harm him: Otsuka, "Killing the Innocent in Self-Defense," pp. 90-91. I think that this is unjustified if the person who is the innocent threat did not take any special risk that led to the conflict.

⁴⁶ Montague, "Self-Defense and Choosing between Lives," p. 211. But compare to the claim that although there should be a distinction between persons who defend themselves (who have a "right" to do so) and third parties (who have an "obligation" to do so), Montague's view is incompatible with this distinction. See Seumas Miller, "Self-Defence and Forcing the Choice between Lives," *Journal of Applied Philosophy* 9, no. 2 (1992): 239-243, 241.

⁴⁷ McMahan, "Self-Defense and the Problem of the Innocent Attacker," p. 261.

all agents, and no substantive line should be drawn between permissions, obligations, rights and duties.⁴⁸

Despite these differences, however, the basis idea underlying the above principles is similar to that underlying *The Principle of Fairness: Responsibility*. It is therefore instructive to consider the criticism made against these principles and its relevance to my thesis.

Several critics reject responsibility as a ground for resolving interpersonal conflicts due to what they consider its "retributive character."⁴⁹ This is a misunderstanding. As I pointed out in the previous part, the justification to sacrifice the responsible person in interpersonal conflicts is not necessarily retributive.⁵⁰ *The Principle of Fairness: Responsibility* in particular rests on the idea that this is the right (fair) way of allocating an unavoidable burden, rather than on retributive ideas.

David Wasserman criticizes Montague's suggestion by claiming that greater fault is neither a sufficient nor a necessary condition for choosing one life over another. First, he claims that greater fault is not sufficient in order to justify self-defense: it is not always justified to sacrifice the person who is most at fault for a forced choice situation. He puts forwards several examples in order to demonstrate this claim. One example involves a choice between two people who will die without a transplant of a liver: the first is not responsible for his predicament, but is old and expected to live only six month even with the new liver, while the second is young and will live thirty years with the new liver, but is responsible for his need of liver due to excessive drinking. Another example requires choosing which of the passengers

⁴⁸ I refer to some of these issues in "Well-Being and Fairness."

⁴⁹ See David Wasserman, "Justifying Self-Defense," *Philosophy and Public Affairs* 16, no. 4 (1987): 356-377, 372; Uniacke, *Permissible Killing*, pp. 185-186; George P. Fletcher, *Basic Concepts of Criminal Law* (New York and Oxford: Oxford University Press, 1998), p. 145 ("linking the culpability of the aggressor with the permissible degree of defensive force seems, however, to confuse the institutions of punishment and self-defense"; although punishment does not necessarily reflect retribution, this remark implies an assumption of retributive punishment since it relates punishment to culpability and contrast punishment with self-defense).

⁵⁰ None of the fault-based principles mentioned above is retributive. See Montague, "The Morality of Self-Defense," p. 88; McMahan, "Self-Defense and the Problem of the Innocent Attacker," pp. 259-260; Draper, "Self-Defense, Collective Obligation, and Noncombatant Liability," p. 66.

of a drowning ship to save, after one of them has ran it into an iceberg, in order to kill another who had raped and caused the death of his daughter. Wasserman does not claim that it is obviously wrong to sacrifice the responsible persons in these examples, but he claims that "even if we are inclined to favor the sacrifice of the culpable parties, it is not the clearly 'dominant' solution that it is in standard self-defense" situations.⁵¹

Wasserman also claims that greater fault is not necessary in order to justify self-defense: it might be justified to sacrifice an aggressor who is not the most responsible person for the attack. For example, he claims that a person who "deliberately incites the aggressor to give himself a legal pretext to kill him," by "relentless insults and harassment," might be "more culpable than the aggressor he provokes," and yet it does not seem that we should help the aggressor if the only options are his death or the death of the provocateur.⁵²

Wasserman is right, I think, in claiming that greater fault, in the sense of responsibility for an interpersonal conflict, is neither a sufficient nor a necessary condition for sacrificing individuals in interpersonal conflicts. But this seems to be compatible with Montague's thesis, and, in any case, it is certainly compatible with my *Principle of Fairness: Responsibility*. This principle, it should be remembered, does not always justify the sacrifice of the person who is most responsible for an interpersonal conflict. Rather, it holds that there is (always) a reason that the person responsible for an interpersonal conflict would bear its cost. I do not suggest that this reason is necessarily decisive; on the contrary, since it might clash with other reasons for action, the resolution of interpersonal conflicts (in which it is applicable) depends on its force and on the force of other, including contrary, reasons for action.

This could be demonstrated with respect to the examples given by Wasserman. I agree that there is room for hesitation with respect to the resolution of all three examples. But this does not undermine *The Principle of Fairness: Responsibility*. On the contrary, in all these

⁵¹ Wasserman, "Justifying Self-Defense," pp. 366-368.

⁵² Wasserman, "Justifying Self-Defense," pp. 368-369.

examples there is a reason that the persons responsible for the situation will be sacrificed. However, this reason is not as strong as in the paradigm case of self-defense,⁵³ and since it sometimes clashes with other reasons, it is not necessarily, and certainly not obviously, decisive. Particularly, the following points should be taken into account concerning the relative force of the competing reasons in these examples.

First, in all examples, the responsibility of the "culpable parties" for the conflict is much weaker than that of a typical aggressor. This is obvious in the ship example, where, we are told, the father was "overcome with rage." It is also true in the transplant example, in which the young person, it is reasonable to assume, did not cause the conflict deliberately but only recklessly. And this is true as well in the provocation example, as provocation is a classical ground for considering persons less responsible for aggression. Thus, in all examples, the reason to sacrifice the responsible person is significantly less strong than in standard self-defense situations. If these diminishing factors were not present, this reason would be as strong as in a typical case of aggression.^{54 55}

Second, in the provocation example it is not at all clear that a person who provokes another, even by "relentless insults and harassment," is more culpable, and especially more responsible for the conflict of lives, than a person who reacts in a way that forces a choice between their lives. In fact, it seems that the contrary is true.⁵⁶

Finally, in the transplant example, there might well be (although this might be disputed) a stronger contrary reason, based on *The Importance Principle*, to prefer the young person, who

⁵³ Gorr asserts, with respect to the ship example, that the "justification for killing those whom Wasserman terms 'past aggressors' is every bit as strong as our justification for killing present aggressor (all else equal)." Gorr, "Private Defense," p. 265. This proposition is too strong as I explain in the text.

⁵⁴ See, with respect to the ship example, Montague, "The Morality of Self-Defense," pp. 82-83.

⁵⁵ Moreover, as pointed out by Gorr, in real life versions of the examples referring to responsibility that is not exhibited at the time the conflict needs to be resolved, there might be doubts with respect to the reliability of the evidence that the persons are indeed responsible for the conflict. See Gorr, "Private Defense," p. 265.

⁵⁶ See Montague, "The Morality of Self-Defense," p. 86.

will live with the new liver thirty years, over the old person, who will live with it only six months.⁵⁷ Thus, there is a contrary reason to prefer the young that affects the overall judgment without casting doubt on *The Principle of Fairness: Responsibility*, namely, on the idea that there is a reason to prefer the old person due to the responsibility of the young for the conflict.

Suzanne Uniacke also claims "that the fact that someone has culpably endangered another person's life does not give me a positive right to sacrifice the culpable party, if necessary, as a means of saving the victim... For example, I have no right to remove a criminal's heart or kidneys on the ground that I put these organs to good use in saving someone else. This is so even if the criminal has, through a prior culpable attack, caused the intended recipient to require transplant." (She acknowledges, however, without explaining the difference, "that in other circumstances, where (say) two people are endangered and both cannot be saved, the fact that one party has culpably endangered the other can be a legitimate ground on which to discriminate against the culpable party.")⁵⁸

This example does not seem to me to undermine *The Principle of Fairness: Responsibility*. There is a natural repulsion from harming persons directly, especially when the harm involves killing them by removing their organs, and this explains why it is easier to accept *The Principle of Fairness: Responsibility* when the choice is who to choose rather than whether to kill actively one person in order to save another. But I do not think that the difference – between killing and letting die – is in itself morally significant.⁵⁹ I think that if, as in the above example, a person is responsible for endangering the life of another and the only way to save the latter involves the sacrifice of the former, it is justified to do so. (Note that the former responsibility for the danger, rather than him being a "criminal" is the relevant consideration.)

⁵⁷ Montague seems to make the same point when he claims that the harm to the young person is more severe than the harm to the old: Montague, "The Morality of Self-Defense," p. 85.

⁵⁸ Uniacke, *Permissible Killing*, pp. 188-189.

⁵⁹ See "Well-Being and Fairness."

Wasserman's own suggestion is that the justification for self-defense is partly based on a morally significant distinction between a person who is responsible for a conflict due to an act that had already been completed at the time the conflict is resolved, and a present aggressor whose responsibility is exhibited at the time the conflict is resolved.⁶⁰ But Wasserman does not offer a persuasive explanation as to why this distinction is morally significant.⁶¹ While the responsibility for the conflict seems of obvious moral significance, the time it is constituted seems morally unimportant in itself. Time might be important only as an indication for responsibility: when a responsibility is exhibited in present aggression, namely, at the time the conflict is resolved, there are typically relatively clear evidence regarding the identity of the responsible person and the nature of his responsibility.⁶² This explains the hesitation from harming "prospective aggressors," pointed out by Wasserman to demonstrate the importance of present aggression.⁶³ (Wasserman also suggests an idea of "deontology of the moment" according to which "a person has significantly greater responsibility for his present than for his past actions."⁶⁴ But this assertion too is left unexplained and seems unpersuasive.)

Seumas Miller criticizes Montague's suggestion on several other grounds. First, he claims that it is based on the mistaken idea "that the attacker by forcing the choice on the defender is somehow responsible for his own death."⁶⁵ I do not think that this is Montague's suggestion, and, in any case, this is not the idea underlying mine. *The Principle of Fairness: Responsibility* is not based on the idea that the responsibility of the aggressor for the conflict relieves the defender from responsibility for his action. Rather, the idea is that the

⁶⁰ Wasserman, "Justifying Self-Defense," pp. 371-377.

⁶¹ See Gorr, "Private Defense," p. 267 ("in the absence of some supporting argument, there is simply no reason to take such an extremely counterintuitive view seriously"). Uniacke accepts the requirement of present aggression, in light of her view that I discuss in the next part, but rejects the arguments offered for it by Wasserman: Uniacke, *Permissible Killing*, p. 185.

⁶² See Gorr, "Private Defense," p. 265.

⁶³ Wasserman, "Justifying Self-Defense," pp. 373-374.

⁶⁴ Wasserman, "Justifying Self-Defense," pp. 375-376.

⁶⁵ Miller, "Self-Defence and Forcing the Choice between Lives," p. 240.

responsibility of the aggressor is a morally significant factor in the resolution of the conflict: it provides a reason to prefer the interests of a non-responsible person to the interests of a person who is.

I thus think that it is a mistake to claim that no justification for self-defense is needed, since the attacker is responsible for a "situation in which someone's life will be lost" but the victim "is not responsible for this situation, it is merely presented to him."⁶⁶ This assertion is misleading. Harming an aggressor is an action that needs to be justified. As I suggested in the previous part, it could and should be justified with reference to the aggressor's responsibility for the existence of the conflict. But, although the aggressor forces a choice, this does not relieve the person who harms him from responsibility for harming the aggressor. Therefore, a justification for this harm needs to be provided.

Miller related criticism on Montague's suggestion is that, while the responsibility of the aggressor for the attack is a morally significant factor, it does not justify self-defense since it is merely another description of the situation.⁶⁷ But obviously the idea reflected in *The Principle of Fairness: Responsibility*, that the responsibility of an aggressor for a conflict between his and another person's interests is a reason to prefer the latter, is a normative idea, rather than only a description of the situation.

Miller also criticizes Montague's suggestion on the ground that in referring only to the choice the aggressor forces it ignores the crucial factor of the aggressor's intention to attack (and harm) another, since a choice could be forced unknowingly.⁶⁸ This point is certainly correct, but it does not apply to my suggestion, which distinguishes between responsible and

⁶⁶ Cheyney C. Ryan, "Self-Defense, Pacifism, and the Possibility of Killing," *Ethics* 93, no. 3 (1983): 508-524, 515-516. Ryan also raises the idea that the mere fact that an attacker *causes* a situation in which a choice needs to be made - even if he caused it innocently - is enough to justify harming him. But, as he himself acknowledges, "it is difficult to imagine what *positive* grounds could be given for this principle." See Ryan, "Self-Defense, Pacifism, and the Possibility of Killing," pp. 516-517.

⁶⁷ Miller, "Self-Defence and Forcing the Choice between Lives," pp. 240-241.

⁶⁸ Miller, "Self-Defence and Forcing the Choice between Lives," p. 241.

non-responsible aggressors (and persons more generally). Moreover, as I noted in the previous part, it is also important to distinguish between degrees of responsibility: causing a conflict intentionally reflects a greater degree of responsibility than doing this negligently. Accordingly, the force of the reason to prefer the non-responsible person varies in accordance to the degree of responsibility.

IV. Other Accounts of Self-defense

In this part, I will continue to explicate the proposed justification, outlined in the previous parts, by comparing it to several other proposals for justifying self-defense and explaining why I find the proposed justification superior to these views.

Before turning to these views, I will note two suggestions that I find unpersuasive on general grounds: the claim that there is an agent-relative permission for agents to prefer their interests to those of others⁶⁹ and deontological claim that self-defense is justified since it does not involve intending harm (or evil) in light of the principle of double effect.⁷⁰ I explained elsewhere why I believe deontological distinctions and agent-relative permissions should be rejected in general,⁷¹ and I find them unpersuasive in this context too.

1. Rights

⁶⁹ For agent-relative conceptions of self-defense see Jeremy Waldron, "Self-Defense: Agent-Neutral and Agent-Relative Accounts," *California Law Review* 88 (2000): 711-750. The idea of an agent-relative justification is also mentioned by Davis with respect to conflicts of innocent parties. She argues that since there is no moral difference between the persons involved in such a conflict, the only basis for preferring one over the other is an agent-relative permission of each to prefer his interests (although the well-being of each is equally important and it is not fairer to prefer one the other). See Davis, "Abortion and Self-Defense," pp. 188-194.

⁷⁰ For an analysis of the applicability of the principle of double effect to self-defense see Shelly Kagan, *The Limits of Morality* (Oxford: Clarendon Press, 1989), pp. 132-144; Uniacke, *Permissible Killing*, chapter 4.

⁷¹ In "Well-Being and Fairness." For a rejection of agent-relative justifications for self-defense on other grounds see McMahan, "Self-Defense and the Problem of the Innocent Attacker," pp. 268-271.

The first kind of proposed justification for self-defense is based on the concept of rights. Reliance on the notion of rights involves three main problems, in my opinion.⁷² First, the notion of rights blur the basic question of the justification of actions, namely, whether actions properly reflect relevant reasons for actions, by focusing on questions such as whether actions violate duties or rights understood as primal concepts, that precede the justification of actions, rather than represent conclusions regarding its outcome. Second, as I have noted in the preface, I think that the notions of rights should be rejected if understood as a fundamental moral concept that deviates from the ideas that individual well-being is the fundamental value that should only be supplemented or, more accurately, adjudicated, in light of consideration of fairness. Finally, a general problem of many ideas which rely on the notion of rights is that the moral significance of the alleged rights is left unexplained: rights-based justifications for self-defense typically assume that, rather than explain why, self-defense is justified. The last problem is exhibited, in particular, in many explications of the justification for self-defense in terms of rights.⁷³

⁷² I explained these points further in "Well-Being and Fairness."

⁷³ For skepticism with respect to the explanatory power of the concept of rights in the context of self-defense see W. A. Parent, "Judith Thomson and the Logic of Rights," *Philosophical Studies* 37, no. 4 (1980): 405-417 ("while the explanatory force of rights per se is problematic, they are nevertheless useful rhetorical devices by which to spur moral argumentation on potentially important issues"); Ryan, "Self-Defense, Pacifism, and the Possibility of Killing," p. 515 ("any argument pro or con which hinges on the issue of rights is likely to get us nowhere"); Wasserman, "Justifying Self-Defense," p. 362 ("even if self-defense can be derived from a right to life, that derivation seems artificial and needlessly oblique"); Gorr, "Private Defense," p. 268 ("for what is 'the right to life' other than a compendium of logically independent rights in terms of which it is explicated?").

Indeed, even Thomson, who persistently explored various rights-based justifications for self-defense, and generally relies extensively on the notion of rights, recognizes (both in general and in this context) the lack of a satisfactory theory of rights and seems to acknowledge that notions of rights simply express normative conclusions. See Thomson, "Rights and Deaths," p. 149 ("I think there does not exist any even remotely plausible theory of the logic of rights"); Judith Jarvis Thomson, "Self-Defense and Rights," in William Parent, ed., *Rights, Restitution, and Risk - Essays in Moral Theory* (Cambridge, Mass.: Harvard University Press, 1986), pp. 47-48 ("It is arguable that if there is to be any point at all in appealing to rights in such discussions, there had better be something independent of permissibilities and impermissibilities which fixes their existence and degree of stringency. It is not obvious that this is true. It might be that to attribute a right is only to talk about permissibilities and

Rights-based justifications for self-defense typically assume that person have rights to life, bodily integrity, property and other interests, which include the right not to be harmed in this respects.⁷⁴ In the context of self-defense, the question is thus why protecting these rights, by defending individuals from attacks, is not forbidden as a violation of the attackers' equivalent rights. There are two possible kinds of answers to this question. The first is that while individuals generally have a right not to be harmed, attackers lack this right. The second answer is that, like people in general, attackers do have a (prima facie) right not to be harmed, so that there is a clash between their rights and the equivalent rights of the people they attack. According to this account, this clash should be resolved in favor of the latter, since their rights are more stringent and thus override those of the former.

One variation of the first kind of rights-based justifications is that persons have only the right not to be harmed unjustifiably and that harming attackers in self-defense do not violate this right as harm in self-defense is justified.⁷⁵ The problem is that this account most obviously begs the question: it clearly assumes that self-defense is justified rather than explain why.⁷⁶

A second variation of the view that aggressors do not have a right to interests such as life, bodily integrity or liberty, and therefore do not have a right not to be harmed, is based on the

impermissibilities, but in a way that groups or collects them, and brings whole clusters of cases to bear on each other... We stand in need for an account of how an appeal to right may be thought to function in ethical thinking."); Judith Jarvis Thomson, "Afterward," in *Rights, Restitution, and Risk*, pp. 253 ("to have a right just *is* its being the case that people may and may not treat you in these and those ways.") and 260 ("we do not yet fully understand how the concept 'has a right to' itself works")

⁷⁴ It was claimed that "the right of self-defense, according to all Natural Rights theories, including Blackstone's, is simply a special case of the right to life": Bedau, "The Right to Life," p. 559. For Blackstone's understanding of the right to life as one component of "the right of personal security," which includes also the protection of "limbs," "body," "health," and "reputation," see Blackstone, *Commentaries on the Laws of England*, book I, chapter 1, p. 129.

⁷⁵ See Judith Jarvis Thomson, "A Defense of Abortion," *Philosophy and Public Affairs* 1, no. 1 (1971): 47-66, 57; Parent, "Judith Thomson and the Logic of Rights," pp. 405-417.

⁷⁶ See Thomson, "Self-Defense and Rights," pp. 37-39; Wasserman, "Justifying Self-Defense," p. 362; Ryan, "Self-Defense, Pacifism, and the Possibility of Killing," p. 513. Indeed, even William Parent - who defends this view (which he calls the "moral specification" view of rights), although only by pointing out flaws in other accounts of rights - acknowledges that (like any theory of rights) it does not offer a justification of actions. See Parent, "Judith Thomson and the Logic of Rights," p. 417.

idea that persons can forfeit this right in light of their conduct.⁷⁷ One suggestion is that individuals forfeit their rights to interests such as life and bodily integrity, and so their right that these interests will not be harmed, by unjustifiably and culpably attacking and endangering others.⁷⁸ A wider suggestion, made by Uniacke, is that individuals forfeit or lose the right to life, and so their right not to be killed, if they pose "an unjust immediate threat to another person's life or proportionate interest." This includes, she suggests, not only culpable aggressors but also blameless aggressors, and even people who do not act at all but pose involuntary or passive threats "provided the threat sufficiently resembles an assault, a clear case being Nozick's man thrown down the well."⁷⁹ (The idea of forfeiture can be suggested either as a complete or as a partial justification for self-defense. The latter view (held by Uniacke) reflects the thought that "the fact that someone does not have a right to life does not itself give me a positive right to inflict lethal force on him or her." On this view, the idea of forfeiture is part of the justification of self-defense, since it removes one moral limit on

⁷⁷ The idea that rights, including the right to life, can be forfeited (though not voluntarily alienated) by a wrongful or criminal act, is mentioned by Locke and Blackstone in various contexts. Locke mentions it, for example, when he writes, while explicating slavery, that a person can "by his own fault, forfeit his own Life, by some act that deserves Death": John Locke, *An Essay Concerning Civil Government in Two Treatises of Government* (1690) (Peter Laslett ed., Cambridge: Cambridge University Press, 1988), section 23 at p. 284. Blackstone similarly writes that life may "frequently forfeited for the breach of those laws of society, which are enforced by the sanction of capital punishment" and that a "malefactor" may forfeit his life: Blackstone, *Commentaries on the Laws of England*, book I, chapter 1, p. 133 and book IV, chapter 14, p. 178. For analysis of Blackstone's view regarding forfeiture and its applicability to self-defense see Bedau, "The Right to Life," pp. 567- 568.

⁷⁸ See Jeffrie G. Murphy, "The Killing of the Innocent," *The Monist* 57, no. 4 (1973): 527-550, 547 ("those attacking have by the attack forfeited certain of their rights"); Joel Feinberg, "Voluntary Euthanasia and the Inalienable Right to Life," *Philosophy and Public Affairs* 7, no. 2 (1978): 93-123, 103-104, 111-112 (who contrasts it with the idea that killing innocent threats, such as "innocent shields," is justified but involves an infringement, though not a violation, of their right to life); and A. J. Ashworth, "Self-Defence and the Right to Life," *Cambridge Law Journal* 34, no. 2 (1975): 282-307, 283, 288-289, 294 (who, however, makes, both the claim that attackers' rights are forfeited and the claim that they are overridden, without distinguishing these claims).

⁷⁹ Uniacke, *Permissible Killing*, pp. 190-209, 229.

harming individuals – their rights – and thus explains why self-defense does not wrong aggressors and is therefore an exception to the prohibition on harming persons.⁸⁰⁾

Another variation of the first kind of rights-based justifications for self-defense, which is closely related, and perhaps equivalent, to the previous ones, is offered by Judith Jarvis Thomson and Uniacke. Both begin with the same assertion that, other things being equal, every person has a right to life, namely, inter alia, a right (against every other person) not to be killed. This right, they continue, generally forbids the killing of persons even if this is the only way to save the lives of others. But, they claim, there is an exception to this rule: persons who attack others or constitute threats (according to Uniacke, "unjust immediate threats") to others - culpably or innocently - violate the right of others not to be killed. Therefore, attackers and threats lose or forfeit their right not to be killed in order to prevent the violation. Therefore, they conclude, it is justified to kill blameworthy as well as innocent aggressors and threats in order to protect the persons who are being attacked or endangered. (This exception does not apply to bystanders who are not related to the attack or threat, although harming them might be justified in light of other exceptions.)⁸¹⁾

(Uniacke considers this variation "as theoretically on a par" with a theory of forfeiture of rights, since "there need be no substantive difference" between them.⁸²⁾ And Thomson says that at least culpable aggressors "forfeit" their right not to be killed and that the crucial common factor is the lack of that right, whether it is through forfeiture or otherwise.⁸³⁾

⁸⁰⁾ Uniacke, *Permissible Killing*, pp. 191-192, 203-204; Draper, "Fairness and Self-Defense," p. 88 ("the fact that an aggressor loses his right not to be harmed is not a reason to harm him – it simply removes one moral obstacle to harming him").

⁸¹⁾ Judith Jarvis Thomson, "Self-Defense," *Philosophy and Public Affairs* 20, no. 4 (1991): 283-310; Uniacke, *Permissible Killing*, pp. 209-227 (Uniacke notes the similarity of her view and Thomson's at footnote 2 at p. 156). In other places, Thomson suggests that it is also justified to kill innocent shields of threat. See Thomson, "Self-Defense and Rights," p. 38; Judith Jarvis Thomson, *The Realm of Rights* (Cambridge, Mass.: Harvard University Press, 1990), pp. 370-371. Compare Uniacke, *Permissible Killing*, pp. 160-172.

⁸²⁾ Uniacke, *Permissible Killing*, pp. 195, 208-209.

⁸³⁾ Thomson, "Self-Defense," p. 301.

Both these variations are problematic. The most obvious concern (which Thomson herself notes in a previous article) is that they seem to imply that, since (both culpable and innocent) aggressors and threats lose their rights not to be harmed, their well-being or interests are of no value and deserve no consideration.⁸⁴ For if the right of individuals not to be harmed (killed) is (at least part of) the basis for the prohibition on harming (killing) them, and if aggressors and threats lack this right, then what is the basis for a prohibition on harming them without limits? Particularly, it thus seems that these views are incompatible with the idea that there need to be good reasons for harming persons, including aggressors and threats, and with the requirements that only necessary and proportional force will be used even against aggressors and threats. For if they have no rights that are (at least part of) the reason not to harm them, it seems that there is no bar from harming them without good reason and more than is necessary and proportional. Thus, the forfeiture variation was criticized on the ground that it leads to the unwarranted conclusions that it is justified to harm aggressors even by agents who are unaware of the attack, and so act without a good reason⁸⁵ and beyond the extent needed to repel the attack (for example, after the attack is over).⁸⁶ As I explained, the idea behind this criticism applies more widely, also to the specification variation and to the requirement of proportionality (which both Thomson and Uniacke endorse, even with respect to culpable

⁸⁴ See Bedau, "The Right to Life," p. 570 ("to say [a person] can forfeit [his right to life] is to say... [he] can justifiably be killed or left to die because now (on account of what he has done) he no longer merits our consideration"). This is indeed the implication of the assertion that since attackers forfeit their rights "there is no conflict" in "interfering with them": Murphy, "The Killing of the Innocent," p. 547. Compare to Kagan's rejection of the suggestion that self-defense against culpable aggressors is justified since it does not involve intending evil on the ground that it permits the employment of unnecessary force: Kagan, *The Limits of Morality*, pp. 136-137.

⁸⁵ Fletcher, "The Right to Life," pp. 144-145.

⁸⁶ Sanford Kadish, "Respect for Life and Regard for Rights in the Criminal Law," *California Law Review* 64 (1976): 871, 884; Thomson, "Self-Defense and Rights," pp. 34-37; Parent, "Judith Thomson and the Logic of Rights," pp. 412-413, 415-416; Joshua Dressler, "Rethinking Heat of Passion: A Defense in Search of a Rationale," *The Journal of Criminal Law and Criminology* 73, no. 2 (1982): 421-470, 454; Wasserman, "Justifying Self-Defense," p. 361; Ryan, "Self-Defense, Pacifism, and the Possibility of Killing," pp. 511-512; Draper, "Fairness and Self-Defense," pp. 87-88.

aggressors).⁸⁷ If aggressors do not have a right not to be harmed (at least during the attack), why is it unjustified to harm them regardless of the relative importance of the competing interests (the harm they might cause and the harm needed in order to prevent them from harming others)?

A suggested answer is that aggressors lose or forfeit their right not to be harmed only to the extent necessary to stop their attack and prevent the danger it poses.⁸⁸ This line of argument might be carried further in order to accommodate also a requirement of proportionality, which is pertinent at the time an aggressor constitutes a risk. It might be claimed that aggressors lose or forfeit their right not to be harmed only (as long as they constitute a threat and) to the extent needed for proportional protection (and retain the right not to be harmed in a non-proportional way).⁸⁹

But these replies seem ad hoc, as they do not explain why the loss or forfeiture is revoked along the lines they suggest. This line of response thus highlights the more general problem of these and other variations that rely on the notion of rights. They do not provide an independent and cogent explication of the basis for their construal of the rights they invoke, that explains the conclusions they draw, and thus it seems that they beg the important questions and are just ways of expressing normative conclusions, rather than justifying such conclusions.⁹⁰

⁸⁷ Thomson, "Self-Defense," pp. 285-286, 301 (it is justified to kill an aggressor in order to save another's life or to prevent a "very grave bodily harm" to another, but not in order to prevent a theft of a "wallet or hat"); Uniacke, *Permissible Killing*, pp. 143-155, 213, 229.

⁸⁸ This answer is suggested by Ashworth, "Self-Defence and the Right to Life," p. 289; and Uniacke, *Permissible Killing*, p. 213.

⁸⁹ Again, this seems to be the assumption underlying the views of Ashworth, "Self-Defence and the Right to Life," pp. 296-297, 303 (the principle of proportionality is entailed by attackers' right to life and "a criminal whose offense merely concerns property has a stronger claim to retain his right to life and physical security than one whose endeavour involves violence"); and Uniacke, *Permissible Killing*, p. 196 (individuals forfeit the right to life if they pose "an unjust immediate threat to another person's life or proportionate interest").

⁹⁰ See McMahan, "Self-Defense and the Problem of the Innocent Attacker," pp. 277-278 (claiming that Thomson's view "is an ingenious exercise in begging the question"; and adding that the above response entails a strange view that determines the scope of rights in light of contingent factors such as

Uniacke acknowledges this point. She confronts this problem by defining the scope of the right to life in terms of what is just (and unjust), as opposed to what is justified (and unjustified) overall. On her view, acts can be just but unjustified, for example, when an agent uses necessary and proportional force, including lethal force, against an unjust aggressor. Such an act does not violate the aggressor's right not to be harmed (killed), but might be unjustified, for instance, if the agent is unaware of the relevant circumstances or acts with a bad motive. Similarly, an action can be unjust (a violation of rights) but justified, for example, harming an unoffending person as the lesser evil.⁹¹

This distinction might indeed avoid circularity - if an independent account of the notion of just (and unjust) and, accordingly, of the right to life, is indeed provided, as a basis for the assumptions regarding the scope of self-defense. And, obviously, such an account will be of value only if it is convincing. However, despite of Uniacke's book-length discussion of "the self-defense justification of homicide" (the subtitle of her book), she writes that such an account "is something that I cannot realistically attempt here," and thus this account "remains important, unfinished business." Moreover, she rejects two kinds of "indirect" moral theories ("a contract theory" and "a two-level theory of optimality," essentially, rule-utilitarianism),⁹² who "have notable contemporary advocated" and "prima facie both seem plausible candidates for generating the required exception." She thus relies only on what she thinks is the "considerable intuitive plausibility" of her assumptions concerning the scope of self-defense.⁹³ Thus, Uniacke suggests only an unrealized possibility of a non-circular theory that supports

the options of defense); Draper, "Fairness and Self-Defense," pp. 86-89 (noting that Thomson does not provide a justification for her view). Compare George P. Fletcher, "Defensive Force as an Act of Rescue," *Social Philosophy and Policy* 7, no. 2 (1990): 170-179, 178. As I have already noted, Thomson herself, despite her extensive reliance on the notion of rights, seems to recognize that reference to rights is only a way of expressing independent normative conclusions.

⁹¹ Uniacke, *Permissible Killing*, pp. 210-218, 230-231.

⁹² Suggested briefly by Kagan, *The Limits of Morality*, p. 135.

⁹³ Uniacke, *Permissible Killing*, pp. 218-228.

her conclusions. In fact, she derives these conclusions in a circular way from intuitions regarding the scope of self-defense.

Furthermore, in my opinion, Uniacke's basic distinction and some of her assumptions regarding the scope of self-defense are unconvincing. First, I believe that there is only one substantive normative standard – an action is right or justified if it best reflects all relevant reasons for actions – rather than two independent fundamental normative standards (just and justified). More specifically, I can see no convincing reason for considering an innocent relation to an attack or a threat as morally significant. Thomson and Uniacke claim that even an innocent (immediate) attack or threat violates the right of the person in danger not to be harmed. But this view is unexplained and, as I explain further below, it is unconvincing, since, by hypothesis, an innocent aggressor or threat is not responsible for the harm that is expected to take place, and this seems to be the only moral factor relevant for the resolution of the conflict.⁹⁴

In another place, Thomson discusses a rights-based justification for self-defense of the second kind I mentioned. She suggests that an aggressor has a right not to be killed, but this right is not absolute and is overridden by the victim's right not to be killed.⁹⁵ Yet, as Thomson herself acknowledges, it is hard to explain how does this idea provide a justification for self-defense (as opposed to an assertion that it is justified).

⁹⁴ For criticism on Thomson's view that innocent aggressors, innocent threats and innocent shields of threats are different from innocent bystanders see Alexander, "Self-Defense, Justification, and Excuse," pp. 53-66; Otsuka, "Killing the Innocent in Self-Defense," pp. 74-94 (innocent aggressors or threats do not violate rights); McMahan, "Self-Defense and the Problem of the Innocent Attacker," pp. 266-268, 275-276 (Thomson's position entails the odd conclusion that justified attackers lose their rights not to be harmed and to defend themselves); Parent, "Judith Thomson and the Logic of Rights," pp. 405-417; Draper, "Fairness and Self-Defense," pp. 86-90. Compare Jerome Hall, *General Principles of Criminal Law* (2nd ed., Indianapolis: Bobbs-Merrill, 1960), footnote 85 at p. 436 (protection against an insane attacker is akin not to defense against a culpable aggressor but to protection from "a natural force," "since an insane person is not bound by duties of the penal law, it follows that he cannot violate any legal right conferred by that law").

⁹⁵ Thomson, "Self-Defense and Rights," pp. 42-48. Compare Bedau, "The Right to Life," p. 569 ("possession" of rights "is not always dispositive of the issue of proper conduct by the individual or treatment by society," since there are other considerations, such as "prior criminal acts" of the right bearer, "scarcity of resources" and the need to resolve conflicts of rights).

Sanford Kadish suggests another right-based justification for self-defense: every person has a moral right against the state "to the law's protection against the (intentional) deadly threats of others" and threats to other important interests (and perhaps against threats to all other lesser interests as well). This right, he continues, includes at least "a legal liberty" to resist such threats "by all necessary means, including killing the aggressor." Kadish offers a contraction reasoning as the basis of this right: "the individual does not surrender his fundamental freedom to preserve himself against aggression by the establishment of the state authority," since such a surrender is only rational when it "yields a quid pro quo of greater, not lesser, protection against aggression that he had before," and this is not the case with respect to self-defense against imminent aggression. Kadish claims that defensive force against aggressors does not violate their right, since they have only the parallel right to resist aggression, which "is not violated by the victim who is only defending against the other's aggression." Kadish applies his analysis both to culpable and to innocent aggressors.⁹⁶

This account, like several other contraction accounts, seems to me unsatisfactory in that it does not explain why persons have a (moral) right to harm aggressors (including innocent aggressors) in defense in the first place. It only states that this right does not disappear after the establishment of a state. Perhaps it is unjustified for a state to prevent persons from defending themselves and others, including by harming (innocent) aggressors, but that in itself does not explain why such harm is justified.⁹⁷

2. The Value of the Interests of Aggressors

⁹⁶ Kadish, "Respect for Life and Regard for Rights," pp. 884-888. Waldron interprets Kadish's view as agent-relative: Waldron, "Self-Defense: Agent-Neutral and Agent-Relative Accounts," p. 747.

⁹⁷ Compare Uniacke, *Permissible Killing*, p. 222.

John Harris claims that self-defense against an unjustified and culpable aggressor is justified (in contexts involving conflicts of lives) since by attacking another, the aggressor "is giving the world a reason to value his life less than that of others. This is because his attack will itself represent an attack on the very value that must support any defense of the value of his own life... An aggressor could not then consistently claim that his own life should be respected while attempting to kill his victim."⁹⁸ Later he adds that "those wicked enough to attempt to inflict such damage on others have... in a sense 'volunteered' for any injury that preventing them may involve."⁹⁹

George Fletcher points to a similar account of the justification of self-defense (which he associates mainly with Anglo-American law). This account is based on "a comparison of the competing interests of the aggressor and the defender... The factor that skews the balancing in favor of the defender is the aggressor's culpability in starting the fight... His interests are discounted, as it were, by the degree of his culpability."¹⁰⁰ This account does not apply, by definition, to innocent aggressors and threats. With respect to culpable aggressors, this account seems to entail a requirement of proportionality, since their interests, though discounted, still receive some consideration. (Proportionality is indeed required by Anglo - American law.¹⁰¹)

⁹⁸ Harris, *The Value of Life*, p. 69.

⁹⁹ Harris, *The Value of Life*, p. 77.

¹⁰⁰ George P. Fletcher, *Rethinking Criminal Law* (Boston: Little, Brown, 1978), pp. 857-858; George P. Fletcher, "Proportionality and the Psychotic Aggressor: A Vignette of Comparative Criminal Theory," *Israel Law Review* 8, no. 3 (1973): 367-390, 377; Fletcher, "The Right to Life," p. 142; George P. Fletcher, *A Crime of Self-Defense: Bernard Goetz and the Law on Trial* (Chicago: The University of Chicago Press, 1988), p. 24. Compare Dressler, "Rethinking Heat of Passion," p. 454 ("self-defense is based on the moral premise that each person has a right to life, and a concomitant right to protect that right, unless and until the original actor unjustifiably violates another person's same or equivalent right. When Victim threatens Actor's right to life, then, the value of Victim's life, in societal terms, is reduced. The harm that flow from his death is negated or, at least, reduced.") It is not clear whether Dressler's refers only to culpable aggressors.

¹⁰¹ See Ashworth, "Self-Defence and the Right to Life," pp. 296-297; Robinson, *Criminal Law Defenses*, volume 1, pp. 86-88; volume 2, pp. 81-88.

The problem is that these accounts do not explain how exactly does an aggression diminish the value of the aggressor's interests. As I explained in the previous part, the responsibility of an aggressor for the conflict is indeed a reason to sacrifice his interests rather than those of another. But this is not because his interests are of no value or are less valuable. The interests of all persons, including culpable aggressors, have value, and accordingly there is a reason not to harm even culpable aggressors. However, the responsibility of aggressors for the conflict is the basis for another - contrary - reason, that they, rather than the persons who are not responsible for the conflict, should bear the inevitable cost of the conflict, namely, the necessity that someone will be harmed. Harris' reference to the inconsistency in the hypothetical position of aggressors is beside the point, especially since their unjustified and culpable attack demonstrates that their judgment is flawed. Harris' remark that aggressors "in a sense" "volunteer" for any injury that preventing them may involve, is similarly puzzling. It is hard to see in what relevant sense this is true, apart from the straightforward normative judgment - not based on any idea of consent - that it is right that the interests of aggressors, rather than those of non-responsible victims, would be sacrificed if some sacrifice in unavoidable.

The suggestion that self-defense is justified due to the diminished value of aggressors' interests is often criticized on other grounds, mainly by claiming that it conflicts with several widely accepted views. Part of this criticism seems to me valid, but occasionally it implies, often implicitly, unwarranted conclusions. One common accusation is that the above suggestion conflicts with the widely held view (assumed to be reflected in law) that the resolution of interpersonal conflicts should be based on the assumption that the lives of all individuals are of equal value. This view ignores, in particular, judgments as to who is morally good or bad (better or worse) person. For example, according to this view, when it is



possible to save only one of two people who would otherwise drown, the fact that one of them is a murderer should not be taken into account in the deliberation whom to save.¹⁰²

This criticism is, in my opinion, correct if interpreted narrowly, as applying only to the assessment of the importance of the lives of the persons involved in the conflict. But this criticism might reflect a wide view according to which the only relevant consideration in the resolution of interpersonal conflicts is the nature of competing interests (I assume that everyone would agree that in a conflict between one person's life and another's property the former should prevail). This stronger interpretation, which is incompatible with *The Principle of Fairness: Responsibility*, should be rejected, in my opinion.

It is also claimed that the above suggestion conflicts with the widely shared view that it is justified to harm aggressors much more severely than the harm they are expected to cause, since it is not plausible to assume that the interests of aggressors should be diminished to a sufficient extent to account for this view.¹⁰³ (One recurring example is the claim that it is justified to kill many aggressors in order to save one person they attack. Again, I am not sure that the number of individuals involved in interpersonal conflicts is in itself morally significant,¹⁰⁴ but obviously this is the assumption underlying this criticism.)

Again, this criticism seems to reflect a valid point. But, in my opinion, it is unwarranted to the extent that implies a rejection of a requirement of proportionality in such conflicts. As I explained, the expected harm involved in any act, including harm to aggressors, is a reason against it, as stated by *The Well-Being Principle*. The force of this reason depends on the magnitude of the act's effects (including the probability that it will cause harm), as pointed out by *The Importance Principle*. Thus, while *The Principle of Fairness: Responsibility* provides

¹⁰² Kadish, "Respect for Life and Regard for Rights," pp. 880-882; Wasserman, "Justifying Self-Defense," pp. 358-359; Rosen, "The Excuse of Self-Defense," p. 47.

¹⁰³ Kadish, "Respect for Life and Regard for Rights in the Criminal Law," p. 882; Wasserman, "Justifying Self-Defense," p. 359; McMahan, "Self-Defense and the Problem of the Innocent Attacker," p. 262.

¹⁰⁴ See "The Relevance of Numbers."

a reason to prefer the non-responsible person, rather than a responsible aggressor, this reason is not necessarily decisive. Despite the responsibility of an aggressor for an interpersonal conflict, it is unjustified, for example, to kill him in order to prevent trivial damage to property, or low probability dangers to minor interests, of (non-responsible) persons.

Harris' idea that the value of the interests of unjustified and culpable aggressors is diminished does not apply with respect to innocent threats and innocent shields of threat. Harris thus concludes that when choosing between innocent parties, there is no reason to prefer an innocent person who is under threat to an innocent threat and an innocent shield of threat. He claims that (in interpersonal conflicts involving one person on each side) we should save one, and it does not matter which, so long as the choice does not "display a vicious preference for one over the other, a preference that effectively treats one as of less value than the other."¹⁰⁵

This account seems to me misguided as well. I agree with Harris that there is no reason to prefer an innocent person who is under attack to an innocent threat or an innocent shield of threat. However, contrary to Harris, I do not think that the choice whom to choose in conflicts between innocent parties is a matter of indifference, as long as the choice does not "display a vicious preference" that "treats one as of less value than the other." Rather, I believe that this choice is important, especially when the competing interests are weighty, such as in a conflict between lives. If there are two contrary competing reasons to protect the well-being of different persons, we have a strong reason, reflected in *The Equal Chance Principle*, to accord each person an equal chance to be preferred. The importance of this reason could be vividly seen by imagining oneself in the place of one of the competing individuals, especially when the interest at stake is important, most notably life. The importance of this reason explains why *The Equal Chance Principle* applies even when there is a contrary reason to prefer one of

¹⁰⁵ Harris, *The Value of Life*, p. 71.

the parties, as long as the contrary reason is not strong enough to cross a certain threshold and outweigh the reason reflected in *The Equal Chance Principle*. For example, in a conflict between two non-responsible persons, *The Equal Chance Principle* should apply even if the interests of one person are more important than those of the other, so that there is a somewhat stronger reason to prefer the former - as long as the difference in importance is not qualitative, in which case the reason to prefer the qualitatively more important interest is strong enough to outweigh the reason to accord each person an equal chance.

3. Unjustified Invasions

The last common rationale for self-defense that I wish to note is based on the idea that an unjustified (wrongful) intrusion by an aggressor, whether responsible or not, on another individual, is morally significant and particularly justifies harming the intruder in order to stop the violation. Many commentators have suggested ideas of this kind as the basis for self-defense.

According to Baruch Brody, self-defense is justified against a person whose "continued existence... poses a threat" to the life of another by "unjustly" attempting to kill him, regardless of whether the person who poses the threat is responsible for this attempt.¹⁰⁶ This claim resembles Uniacke's suggestion that self-defense is justified against persons who pose "unjust immediate threats" whether culpable or not.¹⁰⁷

Fletcher depicts, sometimes sympathetically, an account (which he associates mainly with Kant and German and Soviet law) of self-defense which is based on the idea that unjustified attacks "intrude" or "encroach" upon the "living space," "personal domain," or the "bodily

¹⁰⁶ Baruch Brody, "Thomson on Abortion," *Philosophy and Public Affairs* 1, no. 3 (1972): 335-340, 336.

¹⁰⁷ Uniacke, *Permissible Killing*, pp. 172-193.

integrity" of individuals, and thus violate "individual autonomy" (and, consequently, the "Legal Order"). This account is understood to justify unlimited self-defense as "Right need never yield to Wrong." Particularly, this theory is understood to imply that self-defense is equally justified against any unjustified (wrongful) aggressor, whether culpable or not, without qualification by a requirement of proportionality: "killing an aggressor is permissible if it is the only means available to prevent the invasion of even a minor interest."¹⁰⁸ (However, Fletcher sometimes draws a distinction between the notion of "Right," which ignores the interests of wrongful aggressors, and the notions of "morality" or "justice," which take account of the interests of aggressors and thus supports a requirement of proportionality. He suggests that although the former dominates the law of self-defense, the latter, which he associates mainly with punishment, affect it as well.¹⁰⁹ Fletcher also notes that both German and soviet law, like Anglo-American law, qualify the position that protection of rights in unlimited by introducing a requirement of proportionality, for example, through the doctrine of "abuse of rights."¹¹⁰)

Paul Robinson suggests that the "balance of harms" includes not only "physical harms" but also "intangible harms" such as "the compelling societal interest in preserving the right of "bodily integrity" or "property ownership" and "condemning unjustified aggression."¹¹¹ This

¹⁰⁸ Fletcher, "Proportionality and the Psychotic Aggressor," pp. 378-387; Fletcher, *Rethinking Criminal Law*, pp. 860-874; Fletcher, "The Right to Life," p. 141; Fletcher, *A Crime of Self-Defense*, pp. 32-33, 35-36; George P. Fletcher, "Punishment and Self-Defense," *Law and Philosophy* 8, no. 2 (1989): 201-215, 210, 213-215; Fletcher, "Defensive Force as an Act of Rescue," pp. 171-172; George P. Fletcher, "The Nature of Justification," in Stephen Shute, John Gardner & Jeremy Horder (eds.), *Action and Value in Criminal Law* (Oxford: Clarendon Press, 1993), pp. 175-186, 181; Fletcher, *Basic Concepts of Criminal Law*, pp. 136, 144.

¹⁰⁹ Fletcher, "Punishment and Self-Defense," pp. 208-215.

¹¹⁰ See Fletcher, "Proportionality and the Psychotic Aggressor," pp. 368, 381-387, 390; Fletcher, *Rethinking Criminal Law*, pp. 871-874; Fletcher, "The Nature of Justification," p. 181; Fletcher, *Basic Concepts of Criminal Law*, pp. 136-137, 144-145; Mordechai Kremnitzer, "Proportionality and the Psychotic Aggressor: Another View," *Israel Law Review* 18, no. 2 (1983): 178-214, 210-211.

¹¹¹ Robinson, *Criminal Law Defenses*, volume 2, p. 70.

account seems similar to Thomson's view that self-defense is justified since both culpable and innocent aggression involves violation or infringement of rights.¹¹²

According to Charles Fried, an "unjustified attack creates a relation... of wrong, between the attacker and his victim" and this relation justifies harming the attacker in self-defense. He applies this analysis to intentional attacks, both culpable and innocent (such as that of an insane aggressor) and perhaps even to persons who pose innocent threats without acting intentionally (such as a person who is falling on another).¹¹³

Similarly, Frances Kamm argues that, in conflicts between an innocent threat and another person, the former should generally bear the cost of the conflict, since "it is inappropriate for a person to impose on another person without the latter's permission. To be in that position is to stand in an inappropriate relation to another person (even if the imposition is not anyone's unjust act)."¹¹⁴ (Kamm is not sure whether her thesis applies in all interpersonal conflicts. She writes that it applies at least if the expected harm to both parties is "low but not insignificant" and if the harm to the innocent threat is "minor injury" versus death to the other innocent person.¹¹⁵ She does not explain the grounds for this qualification. Note that, in the last example, when the conflict is between life and "minor injury" (of two individuals none of whom is responsible for the conflict), the thesis I suggest justifies causing the "minor injury," in light of *The Qualitative Difference Principle*, regardless of who constitutes the (innocent threat.)

Michael Gorr does not offer a justification for self-defense (as I noted in the preface, he doubts whether this is possible), but he does suggest some principles for the resolution of

¹¹² Thomson, "Self-Defense," pp. 283-310.

¹¹³ Charles Fried, *Right and Wrong* (Cambridge, Mass.: Harvard University Press, 1978), pp. 44-53.

¹¹⁴ Frances Myrna Kamm, "The Insanity Defense, Innocent Threats, and Limited Alternatives," *Criminal Justice Ethics* 6, no. 1 (1987): 61-76, 65. Kamm suggests that the innocent threat is required to stop the threat and perhaps to compensate the other if the threat has caused him harm.

¹¹⁵ Kamm, "The Insanity Defense, Innocent Threats, and Limited Alternatives," p. 65.

various conflicts, one of which considers threat, including innocent threat, as morally significant. Gorr tries to explain the basis for the intuitive distinction between an innocent bystander - which he thinks it is obviously unjustified to harm in order to prevent a similar harm to another - and an innocent threat - which he thinks it is obviously permitted to harm in order to prevent a similar harm to the victim. He suggests that the fact "that someone poses a threat to another is itself morally significant enough to generate a liberty or permission... to take defensive measures against him but not morally significant enough, in the absence of fault... to trigger a full-blown right to the use of such measures." Gorr acknowledges that "it is difficult to see why the fact that a person plays a casual role in threatening harm to another has any moral relevance whatever" (if this person is not responsible for playing such a role). In fact, he even writes at one point that "the difference between an innocent aggressor and a victim, namely, the fact that the behaviour of the former threatens the latter rather than the other way round, could have the requisite significance only if... [the aggressor] were in some way at fault in bringing about such a threat." And he concludes by saying that the question why it is permissible to harm innocent threats but not innocent bystanders "is a perplexing question for which I do not at present have any good answer." Yet he insists that although he cannot explain why posing an innocent threat is morally significant, it must be so since this is the only way to distinguish between innocent bystanders and innocent threats.¹¹⁶

The main problem with all these accounts is that it is not clear why the fact that one person is being unjustifiably intruded or harmed, justifies harming the intruder, if the latter is not responsible for the situation. The resolution of interpersonal conflicts should take account of the well-being of all persons involved. This is why, as I explained, and contrary to Fletcher's articulation of the "autonomy" view, proportionality is required even with respect to



¹¹⁶ Gorr, "Private Defense," pp. 248, 250, 252-253.

responsible aggressors, despite the reason to prefer non-responsible persons over them.¹¹⁷ When resolving conflicts involving non-responsible persons, the only valid reason to prefer one to the other is the relative importance of the competing interests. This reason, I suggested, is strong enough to outweigh *The Equal Chance Principle* only when there is a qualitative gap in the importance of the competing interests. Otherwise, since the well-being of both innocent parties is important, and fairness does not provide a reason to prefer either based on responsibility for the existence of the conflict, we should follow *The Equal Chance Principle* (assuming that there is no way of avoiding the conflict, for example, by restraining the aggressor without harm). Although an "imposition" or "intrusion" on persons, in the sense of harming them, is no doubt bad, it does not constitute a cogent reason for preferring one non-responsible person to another. None of the above accounts explains why should a person who is not responsible for an unjustified intrusion bear its cost. In other words, it seems that the views reviewed above consider only one side of the equation: certainly there is a reason to protect an innocent party from attack, but there is also a contrary reason to take account of the interests of attackers, especially when they are innocent too.¹¹⁸

¹¹⁷ See part II.

¹¹⁸ Compare to the criticism of Fletcher's view on the ground that the rights of individuals are not "absolute" but rather "context-sensitive," and the resolution of interpersonal conflicts should take account also of other considerations, such as the interests (or "rights") of others and the innocence or culpability of others: Kremnitzer, "Proportionality and the Psychotic Aggressor," pp. 183-189; Larry Alexander, "Justification and Innocent Aggressors," *The Wayne Law Review* 33, no. 4 (1987): 1177-1189, 1182-1184.

See also the rejection of Fletcher's view based on a contractual reasoning or on the idea of equality according to which the interests and rights of aggressors should be taken into account: David A. J. Richards, "Human Rights and the Moral Foundation of the Substantive Criminal Law," *Georgia Law Review* 13, no. 4 (1979): 1395-1446, 1435-1436 (contractual reasoning); David A. J. Richards, "Rights, Resistance, and the Demands of Self-Respect," *Emory Law Journal* 32, no. 2 (1983): 405-435, 426 ("the proportionality requirement rests... on equal concern for both parties' interests in physical integrity. Unjust aggression disturbs the underlying equality in moral rights to basic goods only to the extent required to return the parties to equality").

Indeed, both Fried and Fletcher are aware that their suggestions raise the problem of determining, on non-arbitrary grounds, the difference between defense against an innocent threat and other interpersonal conflicts in which there is a conflict between the interests of innocent parties. See Fried, *Right and Wrong*, p. 53; Fletcher, *Rethinking Criminal Law*, pp. 862-864; Fletcher, *Basic Concepts of Criminal Law*, pp. 143-145.

As Gorr emphasizes, the distinction between an innocent bystander and an innocent threat seems to have an intuitive force. But intuitions should be justified. Moreover, the intuition that a person may defend himself against an innocent attacker could be explained on two grounds. First, intuitions regarding conflicts involving aggression are probably influenced by the fact that in real-life situations most aggressors are culpable, or, at least, assumed to be culpable (since when there is no information regarding this issue, it might be reasonable to assume that the aggressor is culpable on the ground that most aggressors are). Second, the intuition that it is permissible to harm innocent threats in defense could perhaps be attributed to the correct observation that this should be typically excusable, especially in the situations explored by Gorr which involve an immediate and certain danger to life (recall Gorr's qualification that this is "permissible" but not completely right).¹¹⁹

Fletcher rejects the alternative account, that a person who defends himself against an innocent aggressor may be typically excused, on the ground that this view implies two conclusions he considers mistaken. The first is that third parties could not intervene in order to help the victim. And the second is that innocent aggressors who fight back are justified.¹²⁰ The first implication is not disturbing, in my opinion, if properly understood. It is important to remember, in this context, the obvious point that third parties are justified in intervening in order to stop the attack of innocent aggressors and thus resolve the conflict if this is possible without harming either party. Moreover, the thesis I suggest justified intervention by third parties in order to resolve the conflict even if this involves harm to one of the innocent parties, typically the aggressor, if this harm is significantly less severe than the harm that either party is expected to suffer if the conflict is not resolved. Third parties are not justified, however, in

¹¹⁹ For the view that defending oneself against an innocent aggressor or threat is not justified but can often be excused, see Alexander, "Justification and Innocent Aggressors," pp. 1184-1188. Compare Draper, "Fairness and Self-Defense," p. 83 (excuse is especially relevant when people defend themselves against immediate threats).

¹²⁰ Fletcher, "proportionality and the Psychotic Aggressor," p. 375.

preferring the interests of one innocent party to another on other grounds. The second implication is simply mistaken: although a person who defends himself against an innocent aggressor by harming the latter might not be justified, this obviously does not imply that the aggressor is justified. The more plausible assumption is that both parties are only excused.¹²¹

V. Conclusion

I conclude that the analysis suggested in part II seems to me the most plausible account of the justification to resolve interpersonal conflicts by sacrificing the individuals who are responsible for the existence of these conflicts - assuming that other factors are equal. The justification to harm an unjustified and culpable aggressor in the paradigm case of self-defense is the most obvious example. But it is important to bear in mind that this case, conspicuous as it is, is only one example for the applicability of the principles I discussed. These principles might apply in other kinds of interpersonal conflicts, in which the responsibility for the conflict might be very different in nature from that of the unjustified and culpable aggressor in the paradigm case of self-defense, for example, the distribution of health resources, such as scarce organs for transplant. It is also important to remember that the significance of responsibility for the existence of interpersonal conflicts varies according to its degree and the force of the other considerations I mentioned, relating to well-being and to other conceptions of fairness.¹²²

¹²¹ As McMahan rightly notes, if a person who defends himself against an innocent aggressor is not justified but excused, then why does the innocent aggressor who fight back against a similarly innocent (excused) attack is justified? See McMahan, "Self-Defense and the Problem of the Innocent Attacker," pp. 286-287. McMahan criticizes the view that any intrusion justifies harming the intruder on the ground that it seems to entail the odd conclusion that unjustified and responsible (not excused) aggressors are justified in defending themselves against their victims who fight back. See McMahan, "Self-Defense and the Problem of the Innocent Attacker," pp. 257-258 (criticizing what he calls the "Orthodox View" that it is justified to harm anyone who is harming another).

¹²² I discuss several other kinds of interpersonal conflicts in "Well-Being and Fairness."

