1-1-2005

Fairness, Responsibility and Self-Defense

Re'em Segev

Follow this and additional works at: http://digitalcommons.law.scu.edu/lawreview

Part of the Law Commons

Recommended Citation

Available at: http://digitalcommons.law.scu.edu/lawreview/vol45/iss2/3

This Article is brought to you for free and open access by the Journals at Santa Clara Law Digital Commons. It has been accepted for inclusion in Santa Clara Law Review by an authorized administrator of Santa Clara Law Digital Commons. For more information, please contact sculawlibrarian@gmail.com.
FAIRNESS, RESPONSIBILITY AND SELF-DEFENSE

Re’em Segev*

I. INTRODUCTION

The content and scope of normative rules should ideally be derived from a rational analysis of reasons for action and their interaction. However, we often encounter a reverse reasoning, in which the content and scope of norms are assumed to correspond to common intuitive judgments and their justification is explored in the hope of supporting them.1 This reasoning is especially prevalent in the analysis of self-defense (including defense of property and defense of others).2 There is a general agreement, which is rarely found in normative

---

1. This kind of process is part of the method of “reflective equilibrium,” suggested by John Rawls, which explores the justification of “considered judgments” (or, at least, their compatibility to other normative conclusions), as a way of reaching normative conclusions. See JOHN RAWLS, A THEORY OF JUSTICE 17-19, 42-45 (rev. ed., Harvard Univ. Press 1999).

2. This article uses the single term “self-defense,” although my argument is also applicable to defense of property and defense of others. Unfortunately, common terminology concerning defensive force is problematic, especially with respect to the relation of self-defense—in the strict sense—and defense of others. The terms “self-defense” and “private defense” both imply that the justification for defensive force is agent-relative, namely, varies in accordance with the identity of the agent, though in different ways. “Self-defense” implies a distinction between an agent who defends herself and an agent who defends another. “Private defense” implies a fundamental distinction between the employment of defensive force by a private person and by a public official. Although there might be various practical differences between these categories, the assumption underlying the thesis proposed in this article is that all relevant reasons for action and the conclusions they entail regarding the resolution of interpersonal conflicts are agent-neutral, i.e., apply to all agents equally. See discussion infra Part III.
discussions, that self-defense is justified, at least in some paradigm situations. Moreover, self-defense is typically considered the archetype justification for harming individuals and the yardstick for the validity of other stated justifications for harmful conduct. Yet despite this special status of self-defense—or perhaps because of it—its justification is typically presumed rather than explained. When it is explored, many disagreements concerning its foundation and scope emerge and many accounts as to why self-defense is justified seem to be unsatisfactory. Indeed, the "repeated failure of moral and legal theories to establish ... grounding" for self-defense has even led scholars to pessimism "concerning the whole project of attempting to justify the act of defending (innocent) human life" and others to consider self-defense not a justification, but merely an excuse due to lack of sufficient culpability.

3. This agreement is reflected in the law: "all Western legal systems recognize" that self-defense is an exception to the prohibition on intentional killing. George P. Fletcher, The Right to Life, 63 THE MONIST 135, 139 (1980).

4. Both the status of self-defense as the paradigm of justifiable harm and the limited attention to its justification defense are noted by SUZANNE UNIACKE, PERMISSIBLE KILLING: THE SELF-DEFENSE JUSTIFICATION OF HOMICIDE 1 (1994).


The use of defensive force, even deadly force, in protection of self or others (hereinafter "self-defense") is a paradigm of moral and legal permissibility. However, just why this is so has turned out to be a puzzle; indeed, there is today hardly the slightest consensus on the moral justification for this permission to use force, or even on what the various possible alternative justifications might be.

Id.

A similar phenomenon seems to exist concerning the justification of punishment: there is a wide agreement that punishment is justified, at least in some paradigm situations, accompanied by a substantial disagreement regarding the nature and scope of this justification. Douglas Husak argues that a similar phenomenon exists with respect to the doctrine that ignorance or mistake of law is no excuse. See Douglas N. Husak, Ignorance of Law and Duties of Citizenship, 14 LEG. STUD. 105 (1994). However, there is a more significant dissent regarding this doctrine. See Re'em Segev, Justification, Rationality and Mistake: Mistake of Law is No Excuse? It Might Be a Justification? LAW & PHIL. (forthcoming 2005).

6. Michael Gorr, Private Defense, 9 LAW & PHIL. 241, 267-68 (1990) (stating that the justification for self-defense is basic and could not be accounted for in other terms).

7. The distinction between the notions of justification and excuse is complex, and in important respects this distinction is a matter of degree. However, the core distinction between these notions is quite clear and widely accepted. The concept of justification is essentially concerned with whether actions are morally or legally proper, whereas the concept of excuse focuses on whether per-
These pessimistic views are not groundless. Many theo-
sions are sufficiently responsible for actions and particularly on whether persons 
who acted wrongfully or illegally are sufficiently culpable. There are various 
theories of both justification and excuse, but the distinctions between them are 
ot important for the thesis of this article. For discussions that refer to self-
defense as an excuse, see Cathryn Jo Rosen, The Excuse of Self-Defense: Cor-
recting a Historical Accident on Behalf of Battered Women Who Kill, 36 AM. U. 
L. REV. 11 (1986) (claiming that self-defense should be considered an excuse 
rather than a justification). See also Kaufman, supra note 5, at 29 ("The idea of 
justified force as something that society wants unequivocally to approve and en-
courage does not quite fit our intuitive reaction to self-defense, despite modern 
trends."). It should be noted, however, that it is unclear whether Rosen's 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. Id. at 45. The first order interpretation is sup-
ported by the claim that "there probably is no acceptable calculus to support 
treatment of self-defense as a justification." Id. at 49. The second order interpre-
tation is supported by several remarks. One is the assertion that "[t]reatment of 
the battered woman's defense as an excuse does not preclude justifying women 
who kill men under objectively identifiable circumstances more akin to tradi-
tional self-defense." Id. at 44. Another assertion in support of this interpreta-
tion is that "[e]xcused self-defense would better meet the needs of . . . society in 
general." Id. at 45. Finally, there is the assertion that "[f]ew cases in which 
self-defense is claimed . . . fit the model of a justification." Id. at 56. An am-
biguous position, with respect to the nature of self-defense as a justification or 
excuse, is reflected in William Blackstone's assertions concerning the moral 
Note that while Blackstone sometimes employs the terms "justification" and 
"excuse," he does not draw a systematic distinction between justification of ac-
tions and responsibility for actions or culpability. Consider and compare the 
following passages that contain elements of both justification and excuse:

Both the life and limbs of a man are of such high value, in the estima-
tion 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.

1 WILLIAM BLACKSTONE, COMMENTARIES ON THE LAW OF ENGLAND *130 (1860).
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 
足够的 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 ex-
cuse for breaches of the peace, nay, even for homicide itself.

3 WILLIAM BLACKSTONE, COMMENTARIES *3-4.

Homicide in self-defence or se defendendo, upon a sudden array, is also 
excusable, rather than justifiable, by the English law.

4 WILLIAM BLACKSTONE, COMMENTARIES *183. Blackstone's analysis is inter-
preted as reflecting an excuse-based account of self-defense, in light of the last 
passage, by Hugo Bedau, The Right to Life, 52 THE MONIST 550, 559 (1968) and 
Nancy M. Omichinski, Comment, Applying the Theories of Justifiable Homicide 
to Conflicts in the Doctrine of Self-Defense, 33 WAYNE L. REV. 1447, 1449 
ries of self-defense are indeed inadequate in various respects. It is also true that excuse is a proper ground of exculpation in many cases of self-defense. Typically, a person cannot be blamed for defending herself from an attack that might cause her severe harm, regardless of whether the defense is justified. Nevertheless, the aim of this article is to suggest a rational justification for self-defense.

The proposed justification for self-defense is based on a general thesis regarding the proper resolution of interpersonal conflicts, which combines considerations of individual well-being and fairness. This general framework provides an alternative to the moral theories that dominate the inquiry of interpersonal conflicts in general and self-defense in particular: on the one hand, deontological theories and theories of rights, and, on the other hand, aggregative forms of consequentialism, particularly utilitarianism. The interpersonal

---

8. See John Harris, The Value of Life 67-68 (1985); Shelly Kagan, Normative Ethics 90-91 (1998). Jeff McMahan argues that persons who kill innocent bystanders and innocent aggressors in order to save their own lives cannot be excused. His reason is that some cases of this kind, such as that of terrorists, which supposedly should not be excused, are similar to other cases in which there is a threat to life. See Jeff McMahan, Self-Defense and the Problem of the Innocent Attacker, 104 Ethics 252, 287-88 (1994). However, this argument is hasty. The question of whether agents' defense of themselves or others should be excused is a function not of the nature of the threat but of its effects. The nature of threats is only an indication of their effects, and the effects of threats on agents might be different in various cases where innocent bystanders are harmed. Typically, terrorists will be more responsible for their actions, and therefore more culpable, than persons who are caught in unexpected situations in which their life is in danger. For this reason, the latter will often deserve an excuse although the former do not.

9. The general thesis is outlined in Re'em Segev, Well-Being and Fairness, Phil. Stud. (forthcoming 2005); Re'em Segev, Well-Being and Fairness in the Distribution of Scarce Health Resources, 30 J. Med. & Phil. (forthcoming 2005); Re'em Segev, The Concept of Lesser Evil: Beyond Deontology, Rights and Utilitarianism; and Re'em Segev, The Significance of Numbers: Intrinsic and Comprehensive or Instrumental and Restricted? (unpublished manuscripts, on file with the Santa Clara Law Review). In these articles, as in this one, I focus on small-scale, sharp and one-dimensional interpersonal conflicts that involve basic interests such as life, bodily integrity, or property.


11. Influential utilitarian theories include Jeremy Bentham, An Introduction to the Principles of Morals and Legislation (Clarendon Press 1879) (1789); John Stuart Mill, Utilitarianism (2d ed., George Sher,
conflicts usually referred to under the heading of self-defense provide an important test case for central aspects of this framework, and, more specifically, for the interplay between considerations of well-being and one conception of fairness, which is concerned with the moral significance of responsibility of persons for interpersonal conflicts.

The proposed view thus considers the justification for self-defense as entailed by general principles regarding the proper resolution of interpersonal conflict. I believe that situations usually thought of in connection with the notion of self-defense are part of the more general category of interpersonal conflicts, and should therefore be governed by general principles concerning such conflicts, rather than an independent category that should be governed by exclusive principles. In other words, the proposed thesis rests on the assumption that no sharp line distinguishes self-defense cases from other kinds of interpersonal conflicts. To be sure, there are normative differences between various kinds of interpersonal conflicts, but only since part of the general principles are relevant only with respect to certain cases and because the significance and force of the applicable principles might vary in light of the nature of each conflict. Specifically, self-defense situations are unique since the suggested responsibility-based principle apply in special force in some of them, particularly in the paradigm case of self-defense, in which one person unjustifiably and culpably attacks another who in no way contributed to the conflict.

Section II of this article provides a description of the main issues that are considered in it and a general outline of the assumptions underlying the article in regards to the relation between morality and law. Section III outlines the proposed thesis and explains its implications in the context of self-defense. Section IV explores the normative significance of the notion of responsibility, which is central to the proposed thesis. Specifically, Section IV compares the suggested responsibility-based principle to other principles that are based on a similar notion of responsibility and examines vari-
ous objections made against these principles.\textsuperscript{14} Finally, Section V explains why I believe the proposed thesis is preferable to other kinds of dominant accounts of self-defense.\textsuperscript{15}

II. THE QUESTIONS

The article explores three main issues. The first and most fundamental question is the nature of the justification for the paradigm case of self-defense, namely, the justification for harming an aggressor who unjustifiably and culpably attacks and endangers the well-being of another who is not responsible for the attack and thus for the existence of the conflict. According to the proposed thesis, self-defense situations should be viewed as a special kind of interpersonal conflicts, since they require a choice between the well-being of the aggressor and that of the person under attack (assuming that harm to at least one of them is unavoidable). According to this thesis, the justification for self-defense is therefore based upon a set of principles concerning the resolution of interpersonal conflicts in general, which identify several morally significant considerations, based on the notions of well-being and fairness, and resolve clashes between these considerations.

The second issue is the implication of the basic justification for self-defense with respect to a requirement of proportionality in the employment of defensive force. There is wide agreement that it is justified to harm an aggressor in the paradigm case of self-defense if this is necessary in order to prevent the aggressor from harming an innocent victim, even if the conjectured harm to the aggressor is more severe than the conjectured harm the attack creates. For example, it is generally considered justified to kill an aggressor in order to prevent a severe, but not deadly, bodily injury to the attacked person.\textsuperscript{16} However, there is a dispute as to whether the last proposition should be qualified by a requirement of propor-

\textsuperscript{14} See discussion infra Part IV.
\textsuperscript{15} See discussion infra Part V.
\textsuperscript{16} The standard example used to illustrate this point is of a person who defends her life by killing several aggressors. I avoid 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 Segev, The Significance of Numbers, supra note 9. In order to bypass this complex issue, this article focuses on conflicts involving only one person on each side of the conflict.
tionality. Such a requirement would not justify severely harming an aggressor in order to prevent a trivial danger to her victim, for example, killing an aggressor in order to keep her from damaging a trivial property of another. The proposed thesis implies a requirement of proportionality, which excludes such actions.

The third issue is the normative significance of responsibility of individuals for the existence of interpersonal conflicts in determining the proper resolution of such conflicts. The suggestion in this regard is that culpable aggressors and others who are responsible for the existence of interpersonal conflicts, based on a choice-based conception of responsibility, should be distinguished from other persons involved in interpersonal conflicts who are not responsible for their existence. (Throughout the article, I will use the term "responsibility" in the sense of a choice-based conception of responsibility, unless I explicitly note otherwise.) The category of responsible agents must be broken down into sub-categories since individuals might be responsible for interpersonal conflicts to various degrees, and the degree of responsibility affects its normative significance. The category of non-responsible individuals includes, inter alia, the following common sub-classifications: (1) "innocent aggressors": individuals who en-

17. The necessity requirement, namely, the requirement that the force used will be necessary in order to stop the attack, and the requirement of proportionality, are independent of each other. Suzanne Uniacke argues that "[a]lthough necessity and proportionality are distinguishable conditions . . . considerations of proportionality . . . must sometimes form part of the normative background against which necessity is judged." See UNIACKE, supra note 4, at 83. However, such a sweeping claim, that the requirements of necessity and proportionality are necessarily interconnected with respect to every conception of necessity and every conception of proportionality, seems misguided. Indeed, the contended interrelation between these requirements is not even demonstrated in the examples provided by Uniacke. Generally, these examples involve cases in which force is strictly necessary, regardless of considerations of proportion. Her first kind of example provides that, "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." Id. However, in this case, 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. Her second kind of example provides that the endangered agent could have avoided being caught in the dangerous situation. See id. This case does not, however, demonstrate that force is not strictly necessary, but rather, as explained in the next section, that other considerations, concerning prior responsibility, might also be relevant in the resolution of interpersonal conflicts.
danger others knowingly, even intentionally, but are not responsible for their behavior (for example, insane aggressors); (2) "innocent threats": individuals whose body constitute physical threats to others without their control (for example, a person who is thrown by another down a well at the bottom of which there is a third person who will be crushed and killed by the body of the falling person); (3) "innocent shields of threats": individuals 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 and halted without also harming her); and, most obviously, (4) "bystanders": individuals who are not related to the danger in any way but sacrificing or actively harming them is the only way to save others (for example, each of two drowning persons only one of whom could be saved, or a person whose property could be harmed in order to save the property of another person). The proposed view is that since all these subgroups include individuals who are not responsible for the existence of the relevant conflict, according to a choice-based conception of responsibility, they should be treated in the same manner. The proposed thesis thus differs from the common view that considers causal responsibility, for example, the responsibility of innocent threats, as a morally significant factor. Accordingly, the term "innocent" in this article refers to individuals who are not responsible, according to a choice-based conception of responsibility, for the existence of the pertinent interpersonal conflict. This use should be distinguished from the common use of the term "innocence," which refers to a causal conception of responsibility and thus considers as responsible, for example, innocent threats.

There are, of course, other aspects of the justification and

18. The terms and quotes concerning the last two categories are from ROBERT NOZICK, ANARCHY, STATE, AND UTOPIA 34-35 (1974) (suggesting that a self-defense justification is applicable to innocent threats and leaves open the question of innocent shields of threats).

19. See Kaufman, supra note 5, at 24 ("it is almost universally accepted that one has a right to use deadly force against even an innocent aggressor, such as the psychotic aggressor or the mistaken aggressor").

20. See, e.g., Thomas Nagel, War and Massacre 1 PHIL. & PUB. AFF. 123, 139 (1972); Judith Jarvis Thomson, Rights and Deaths, 2 PHIL. & PUB. AFF. 146, 154 (1972); UNIACKE, supra note 4, at 78-81, 95-96.
content of self-defense. However, I focus on the above three issues since they are the ones which are most directly relevant to the main aspects of the thesis I suggest.

The analysis of the justification for self-defense, and its boundaries, essentially applies to both moral and legal norms. The law should not necessarily reflect every valid moral standard, since some moral standards would 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. Otherwise, the law should reflect proper moral standards. This is generally the case in the context of interpersonal conflicts involving basic interests that are threatened by serious adverse effects, since, in these cases, the above considerations are either not relevant or, when they are relevant, are not decisive. Indeed, although there is a disagreement as to the proper limits of the law, it is widely accepted that it should at least be concerned with the protection of important interests of individuals from harm. While this idea involves various difficult questions

21. Examples for other questions are whether, within a requirement of necessity, safe retreat should be generally required and 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 intent to defend herself or someone else. Since the answer to the last question might affect my analysis at several points, I should note that I assume that a rational belief in the existence of the relevant justificatory circumstances is a precondition for justification. I explain why in Segev, supra note 5. This issue is discussed in the context of self-defense by, for example, Russell Christopher, Self-Defense and Defense of Others, 27 Phil. & Pub. Aff. 123 (1998).

22. A parallel condition might be relevant also in the moral sphere, so that only reasons beyond a certain threshold of importance should be the basis for moral norms. See Shelly Kagan, The Limits of Morality 68 (Clarendon Press 1989). Obviously, the threshold in the legal sphere should be higher in light of the additional cost of legal regulation.

23. There are various ways of reflecting normative standards. In this context we should bear in mind the distinction between ultimate moral standards and the public—including legal—rules that would best further the aims of the ultimate standards. Occasionally, an exact reflection of ultimate standards in public rules would not be the optimal method of furthering the aims of these standards. Particularly, there are often strong (though not necessarily decisive) reasons for formulating public—including legal—rules in advance and in a relatively clear form. This might occasionally justify, for example, less precision at the margins for the sake of clarity.

24. The classic 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 through the law, is by John Stuart Mill, On Liberty (David Bromwich et al. eds., Yale Univ. Press 2003) (1859).
about its proper content and boundaries, its core seems both justified and widely agreed upon, and the main disputes concern the question of whether the law should go beyond this sphere. 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.

III. FAIRNESS AS THE BASIS OF SELF-DEFENSE

A. The General Thesis

The general thesis I suggest for the resolution of inter-


26. For discussions of the legitimacy of other grounds for invoking the law, see H. L. A. HART, LAW, LIBERTY AND MORALITY (Oxford Univ. Press 1963); 1 JOEL FEINBERG, HARM TO OTHERS (1984); 2 JOEL FEINBERG, OFFENSE TO OTHERS (1985); 3 JOEL FEINBERG, HARM TO SELF (1986); 4 JOEL FEINBERG, HARMLESS WRONGDOING (1988).

27. While emphasizing the connection between moral and legal standards, inter alia, in the context of self-defense, Kent Greenawalt suggests that there are distinctions between the moral and the legal spheres beyond those suggested above (he refers mainly to prevailing legal norms, but his discussion clearly shows that he also considers proper moral and legal standards to be closely connected). For example, he states that even if there is a moral distinction between culpable and innocent aggressors, it should not necessarily be reflected in the law as 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.” Kent Greenawalt, Violence - Legal Justification and Moral Appraisal, 32 EMORY L.J. 437, 474 (1983). 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.” Id. Greenawalt’s factual observation may well be correct, and it should affect the evaluation of the actions of agents who protected themselves or others from attackers that they reasonably but mistakenly assumed to be culpable. But I fail to see how this empirical observation, in itself, justifies a difference between moral and legal standards in this context. In principle, this observation is relevant to the evaluation of the justification of actions both in the moral and in the legal spheres. If there is a distinction between culpable and innocent aggressors, and if the assumptions of agents in this respect are relevant for the evaluation of the justification of their actions, as Greenawalt rightly assumes, why not reflect this distinction not only in morality but also in the law, even if it will be relevant only in rare cases? There might be various evidentiary considerations in this regard that are especially relevant in institutional, and, particularly, legal settings, but such considerations should not affect cases in which the relevant factual factors are not in disputed.
Justification for Self-Defense

Personal conflicts rest on two basic ideas. First, the fundamental value and source of reasons for actions, in the context of interpersonal conflicts, is individual well-being. Second, interpersonal conflicts—in which reasons for action to protect the well-being of different persons clash—should be resolved in light of several conceptions of fairness, which reflect the independent value of persons, the difference in the importance of various aspects of individual well-being and the moral significance of responsibility of persons for interpersonal conflicts.

In other words, according to the proposed theory, individual well-being is what we should fundamentally care about whereas the concept of fairness is concerned with how we should care for this value when we are confronted by interpersonal conflicts.

It is important to emphasize that, according to the proposed thesis, the concepts of individual well-being and fairness do not represent competing fundamental values: the idea that interpersonal conflicts should be resolved fairly completes the idea that individual well-being is the fundamental value, rather than constitutes a contrasting idea concerning the fundamental value. This feature becomes clear once we notice that the concept of fairness becomes applicable only when there is an internal clash within the concept of individual well-being.

These two basic ideas are elaborated in five principles that reflect agent-neutral (rather than agent-relative) and impartial reasons for action:

1. The Well-Being Principle reflects the intrinsic value of each person. It holds that there is a reason to protect, enhance, and at the least to not adversely affect the well-being of persons.

28. For an elaboration of these ideas and the principles they entail, see Segev, Well-Being and Fairness, supra note 9 and Segev, The Concept of Lesser Evil, supra note 9.

29. The agent-neutrality and particularly the impartiality assumption could be stated more accurately in the form of a preliminary principle. The Impartiality Principle reflects the equal value of each person. It holds that reasons for actions are agent-neutral, that is, apply equally to all agents and particularly that reasons for actions should be evaluated from an impartial perspective, instead of an agent-relative perspective that accords special weight to personal aspects of agents' lives. See Segev, The Concept of Lesser Evil, supra note 9.

30. Segev, Well-Being and Fairness, supra note 9; Segev, The Concept of Lesser Evil, supra note 9.
(2) The Equal Chance Principle, which reflects a consideration of fairness, addresses clashes that arise within the Well-Being Principle between conflicting reasons to protect the well-being of different individuals. It is based upon the following considerations: (a) each person has an *intrinsic and independent* value and therefore the fate of each person is intrinsically and independently significant; (b) 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 (c) there is no common denominator with regard to which interests of different persons could be traded-off *without a loss*. In light of these truisms, the Equal Chance Principle, as an aspect of the more general 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 reason should be elaborated in the form of the following two hierarchical sub-principles: (a) there is a reason to distribute benefits or inevitable costs between all persons equally, so that each would get the maximum possible equal benefit or bear the minimum possible equal loss, or, when exact equality is impossible, so that each person would get a roughly equal benefit or incur a roughly equal loss, provided that the benefit or reduction of cost for each person is significant; (b) if it is impossible to save or benefit each person to a significant and roughly equal degree, the second best option is to accord each person the highest possible *equal chance* to be preferred, and then save, or otherwise benefit, the person chosen in this way. For simplicity, this article assumes that it is impossible to prevent harm with respect to each person to a significant and roughly equal degree and for this reason the second best option applies and the Equal Chance Principle bears a literal meaning.

(3) The Importance Principle, which reflects another conception of fairness, signifies the difference in the importance of various aspects of individual well-being. It holds that the strength of the reason provided by the Well-Being Principle depends on the importance of the interest at stake, the prob-

---

31. This point is elaborated in Segev, *Well-Being and Fairness in the Distribution of Scarce Health Resources*, supra note 9.
ability that it will be affected in case of inaction, and the probability that an action will affect it. The Importance Principle thus incorporates and elaborates the Well-Being Principle. If the two probabilities above are equal, the more important is the interest, the stronger is the reason to protect, enhance and not harm it. Therefore, in resolving interpersonal conflicts, there is a reason to prefer the person who would be affected most significantly by the resolution of the conflict (e.g., there is a reason to prefer the person whose life could be saved over a person whose property could be saved).\textsuperscript{33}

(4) The Substantial Difference Principle resolves the potential clash between the Equal Chance Principle and the Importance Principle. It holds that the Importance Principle prevails over the Equal Chance Principle if (taking relevant probabilities into account) there is a substantial gap in the importance of the competing interests.\textsuperscript{34}

(5) The Principle of Fairness-Responsibility, which may be considered as a third aspect of the concept of fairness, accounts for the rational power of persons by making them accountable for their choices. It holds that, when an interpersonal conflict requires a choice between the well-being of individuals, there is a reason to prefer those who are not responsible for the existence of the conflict over those who are, and those who are less responsible over those who are more responsible (according to a choice-based conception of responsibility). In other words, the responsibility of a person for an interpersonal conflict in which at least one person must bear a setback to her well-being or forgo a benefit is a reason that the responsible, or the more responsible, person, should bear the burden or enjoy the benefit involved in the resolution of the conflict, rather than the non-responsible, or the less responsible, person.\textsuperscript{35}

The resolution of potential clashes between the Principle of Fairness-Responsibility and the previous principles depends on the relative force of the conflicting principles, which is determined in light of the importance of the competing interests, the relevant probabilities, and the degree of responsibility of each person for the existence of the conflict.

\textsuperscript{33} Id.
\textsuperscript{34} Id.
\textsuperscript{35} Segev, The Concept of Lesser Evil, supra note 9.
The proposed set of principles does not provide specific conclusions regarding the resolution of each of the wide variety of potential interpersonal conflicts. The aim of these principles is more modest: to identify the main normative factors that are relevant in the context of interpersonal conflicts, including self-defense settings, as well as the major possible clashes between them, and to point out the general outline of the way in which these clashes should be resolved.

The proposed principles are different from two kinds of influential moral theories in several respects. First, the proposed principles differ from standard deontological theories and theories of rights, which reject the basic consequential idea that it is always right (and required) to act in the way that is expected to lead to the best possible state of affairs, and embrace agent-relative constraints (framed in the language of duties or rights) that forbid certain actions, particularly those that positively harm individuals, even when these actions lead to the best possible consequences. The theory presented in this article is different from these theories in two respects. First, the proposed framework considers the basic normative unit to be reasons for action and the right action as the one that best reflects all valid reasons for action. In contrast, consequential theories consider the basic normative unit to be states of affairs and the right action as the one that leads to the best possible overall state of affairs. Second, the proposed principles reflect agent-neutral, rather than agent-relative, reasons for action. Third, the suggested principles reflect the idea that individual well-being is the fundamental value (rather that non-consequential duties or rights). According to the proposed thesis, the value of individual well-being needs to be construed, supplemented or adjudicated in interpersonal conflicts, in light of the above conceptions of fairness, in order to convey proper concern to the well-being of different persons when resolving interpersonal conflicts. However, unlike deontological and rights-based theories, the proposed principles, and particularly the concept of fairness which they reflect, do not substitute or limit the fundamental value of individual well-being in light of other fundamental values, such as non-derivative duties and rights. Despite the intuitive appeal of deontological and rights-based theories, and their prevalence in current philosophical discourse, it is hard to justify them and therefore I suggest that they should
be rejected.\textsuperscript{36}

The proposed theory also differs from standard consequential theories, and particularly from utilitarianism, in various respects. First, it focuses on reasons for action and the best way of reflecting them, rather than on states of affairs and the best way of maximizing various features of them. Second, it relies on conceptions of fairness, and consequently on distributive and backward-looking considerations. Third, it rejects sweeping aggregation of different kinds of interests, regardless of their nature, importance, or effect on persons as individuals. The last two points are, of course, familiar criticisms of utilitarianism. As many have noted, aggregative forms of consequentialism, particularly utilitarianism, do not properly reflect the independent value of each person as an individual.\textsuperscript{37}

Despite these differences, the proposed theory is nevertheless essentially consequential in one important respect: it considers individual well-being as the \textit{fundamental} value and source of reasons for actions. The other suggested reasons for action do not represent competing fundamental values, but rather reflect what is considered the proper or fair way of resolving interpersonal conflicts of well-being. Moreover, the proposed theory goes as far as any complete moral theory conceptually could and, in my opinion, substantively should, in endorsing the idea that the fundamental standard for the rightness of actions is their effect on individual well-being. Conceptually, every complete normative theory - that is, a theory that aims at providing general guidance for actions -


\textsuperscript{37} This criticism is widely accepted since Rawls, \textit{supra} note 1, at 22-30, 66-69, 175-92. A more complicated question, which is beyond the scope of this article, is whether what is wrong with utilitarianism is aggregation itself, the sweeping nature of utilitarian aggregation, or merely its indifference to distributive and fairness-based considerations within an aggregative framework. The latter view seems to be the more prevalent. For this view see Derek Parfit, \textit{Justifiability to Each Person}, 16 RATIO 368, 379-84, 389 (2003). But I believe that aggregation itself raises difficulties. I explore the complex question of the moral significance of various forms of aggregation in Segev, \textit{The Significance of Numbers}, \textit{supra} note 9.
must go beyond this idea in order to resolve internal clashes within it, such as those that emerge in interpersonal conflicts of well-being. Indeed, even utilitarianism—the paradigm of a consequential theory—does not and cannot, as a complete moral theory, simply declare that a certain good, like well-being, is the fundamental normative standard. Its governing principle is rather the maximization of the aggregate balance of well-being. The additional normative content of this further step is often overlooked or trivialized, perhaps due to the common assumption that the only rational way to convey concern for a certain value is by adopting an aggregative standard. From a substantive point of view, however, I believe that this assumption is mistaken: an aggregative standard is irrational since it does not properly reflect the independent value of persons. Thus, the suggestion is that the idea that a certain value, such as, individual well-being, is morally important both could and should be separated from the aggregative standard that typically accompanies it.

The common contrast between utilitarianism and deontological theories or theories of rights overlooks the possibility of a theory of the kind proposed in this article, which considers individual well-being as the fundamental value, but resolves interpersonal conflicts of well-being in light of non-aggregative conceptions of fairness.

This article does not attempt to defend the proposed principles in their general form. Rather, its aim is to support these principles by demonstrating that they offer a plausible justification for self-defense. In light of this goal, the article will not explore in depth the aspects of the proposed theory that are not directly relevant to the justification for self-defense. It is important, however, to clarify several major aspects of the principle that play a crucial role in the exploration of the justification of self-defense within the framework of the proposed set of principles - the Principle of Fairness-Responsibility. The following brief remarks, in the next section, do not exhaust the complex and wide issue of responsibility, but only note some of the basic assumptions regarding the nature and scope of the conception of responsibility that underlines the proposed framework.

B. The Principle of Fairness-Responsibility

According to the Principle of Fairness-Responsibility, in
the resolution of interpersonal conflicts, there is a reason to prefer individuals who are not responsible or are less responsible for the existence of a conflict in which they are involved to those who are responsible or more responsible, respectively. This principle rests on two related ideas. First, the capacity of persons for rationality is a ground for holding them accountable for their choices. Second, based on the significance of individual autonomy, as an important element of individual well-being, there is a reason to confer a benefit to the person who is responsible for its existence, rather to another person, and to make a person who caused a burden to bear it, if someone must.

The Principle of Fairness-Responsibility reflects a conception of responsibility that is based on choice, rather than on causal responsibility. This conception reflects the assumption that random facts should not affect normative conclusions. Therefore, according to this conception, if a person constitutes an innocent threat or an innocent shield of threat not due to a choice she made, then she should be treated as a bystander.

It is important to emphasize that the Principle of Fairness-Responsibility does not reflect a retributive idea, accord-

---

38. For an extensive discussion of the idea that people should be held responsible only for their free choices, insofar as they affect others, see ERIC RAKOWSKI, EQUAL JUSTICE (Clarendon Press 1991).

39. This conclusion is reinforced by the fact that typically persons who are responsible for benefits invest resources in order to create them, and persons who are responsible for burdens benefit from the choices that led to them.

ing to which there is a reason to give a reward for a right action and a penalty for a wrong action.\textsuperscript{41} Rather, this Principle holds that when a choice between the interests of several individuals is inevitable, namely, when there is an interpersonal conflict, there is a reason why the person who is responsible for the conflict's existence should bear the cost instead of a non-responsible person.\textsuperscript{42}

The Principle of Fairness-Responsibility is thus based on a non-retributive, choice-based conception of responsibility. Therefore, this principle applies both to wrongful and to justified choices and actions. According to the suggested choice-based conception of responsibility, a person is responsible for the existence of a conflict if she chooses to perform an action that led to it. Generally, this principle applies when a person chooses to take a risk that might lead to a conflict between her interests and those of others. For example, the Principle of Fairness-Responsibility applies when a person chooses to engage in a potentially risky recreational activity, such as swimming, and is caught in a storm so that the only way to save her life requires substantially damaging a boat. If it is possible to use either the swimmer's boat or another person's boat, the Principle of Fairness-Responsibility holds that, other things being equal, we should use the swimmer's boat.\textsuperscript{43}

The Principle of Fairness-Responsibility's reliance on a choice-based conception of responsibility should be contrasted with a causal conception of responsibility. This principle does not apply to persons involved in an interpersonal conflict if they are not responsible for a choice that led to it, even if they constitute or are a part of a threat. This group includes, inter alia, innocent aggressors, innocent threats, innocent shields of threats, and bystanders.\textsuperscript{44} While there is a reason to prefer a non-responsible person to a responsible person when resolving an interpersonal conflict, there is no reason to prefer one person to another just because she causally constitutes a

\textsuperscript{41} The difference between the notions of fairness and retribution, in the above senses, is noted, for example, in KAGAN, \textit{NORMATIVE ETHICS}, \textit{supra} note 8, at 55-57.

\textsuperscript{42} This is why I prefer to call this the "Principle of Fairness-Responsibility" rather than just the "Principle of Responsibility."

\textsuperscript{43} In addition, if the only available boat belongs to another person, the Principle of Fairness-Responsibility holds that there is a reason that the swimmer should compensate the boat's owner for her loss.

\textsuperscript{44} For an explanation of these terms, see discussion \textit{supra} Part II.
threat or a part of a threat. In other words, there is no reason to distinguish between non-responsible or innocent persons, whatever their role in the conflict: bystanders, (innocent) aggressors, (innocent) threats, or (innocent) shields of threats. Indeed, since there is no substantive difference between various categories of innocent (non-responsible) persons, it becomes hard to draw the lines between these categories on non-arbitrary grounds.\(^4\)

On the other hand, the non-retributive character of the Principle of Fairness-Responsibility makes it applicable beyond agents who are responsible for wrongful actions, such as an aggressor who unjustifiably and culpably attacks another and thereby creates a situation in which one of them must be harmed. This principle applies also to agents who choose to perform justified actions that create a special risk of generating an interpersonal conflict.\(^4\) For example, the decision of a person, such as the swimmer in the previous example, to engage in a recreational activity that involves a slight risk of leading to an interpersonal conflict (which requires a choice between her interests and those of another). A decision to engage in such an activity might be justified, due to its benefits, including the pleasure the agent derives from it, despite the slight risk it involves. Nevertheless, under the proposed conception, this person might be responsible for the existence of a

---

\(^{45}\) For the implication of the proposed theory with respect to defense against non-responsible aggressors and threats see discussion infra Part III.C. Remember also that even when it is unjustified to harm one innocent person in order to save another, including when the former is an innocent aggressor, threat or shield of threat, it does not necessarily follow that the defender is blameworthy for doing so. He might be excused. See discussion supra Part I.

\(^{46}\) For a similar argument, see ROBERT M. VEATCH, TRANSPLANTATION ETHICS 315 (2000) ("The critical variable is not whether the behavior is moral or immoral but whether one could have knowledge of the risks involved and voluntarily choose whether to engage in the behavior . . . . [T]here could be a behavior that we do not consider ethically suspect in the slightest but nevertheless poses a significant risk for needing an organ."). There is, however, one exception to the rule that the Principle of Fairness-Responsibility is concerned with the responsibility for instead of the justifiability of actions: persons who take justifiable risks in order to protect the well-being of themselves or of others. For example, the fact that a firefighter is fully responsible for a risk of harming herself during her professional activity does not make her less eligible for help if the risk materializes. See id. at 316; Jules Coleman & Arthur Ripstein, Mischief and Misfortune, 41 McGill L.J. 91, 111 (1995) ("[R]escuers are allowed to expose themselves to risks that would normally preclude recovery from their injurers . . . . [T]he special treatment accorded rescuers supposes their activity to be important enough to make them bear proportionally less risk.").
conflict that is attributed to the risks she created. If this is indeed the case, the Principle of Fairness-Responsibility suggests that she should bear the cost involved in resolving the conflict. Particularly, the Principle of Fairness-Responsibility applies whenever a person made a choice to take a risk that could lead to a conflict between her interests and those of others, regardless of whether she was aware of this consequence or whether she aimed to create such a conflict.47

The Principle of Fairness-Responsibility refers to responsibility for dangers, and consequently for interpersonal conflicts, rather than to the temporal or spatial proximity of the responsible person to the agent. For this reason, the Principle covers persons who are responsible for the existence of interpersonal conflicts even if they are not the direct or immediate source of the danger or are not in the danger’s vicinity. 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 at a later time or at a distance place (relative to the time and place in which the responsible party performed the relevant action). In other words, the Principle of Fairness-Responsibility is applicable not only to responsible aggressors, but also to what might be called responsible “bystanders,” i.e., persons who are not a current or direct threat but have made a choice that led to an interpersonal conflict.48 Whether we refer to the justification for harming such persons under the category of self-defense, or under the broader category of interpersonal conflicts, is of course immaterial.

People may be responsible for various kinds of actions and inactions, as well as for conflicts that might ensue from them. Consequently, the Principle of Fairness-Responsibility has wide and diverse implications in many contexts. For example, it provides a reason to give the benefit that resulted from the work of one person to the person who performed the work rather than to another who chose to perform some leisurely activity instead of working. The practical implications

47. The knowledge and aim of the agent might, however, affect her degree of responsibility.

48. Sometimes it will be practically impossible to resolve the conflict by harming a responsible person who is not present at the vicinity of the conflict, but occasionally this might be possible.
of this principle might, however, be limited in many contexts in which individuals cannot be held responsible for various situations or can only be held responsible for situations to a limited degree. Nevertheless, it seems that individuals can be held responsible for at least small-scale and one-dimensional interpersonal conflicts that are the result of relatively clear choices, such as those described above in the examples of the reckless swimmer or the culpable aggressor.  

Responsibility is a matter of degree, particularly for interpersonal conflicts. Under the preferred conception of responsibility, the extent of the responsibility of a person for an act and its consequences appears to be dependent mainly on the nature and degree of the risk the act is expected to create. For example, although a person who engages in a recreational activity that involves a slight risk of creating an interpersonal conflict might be responsible for such a conflict, she is less responsible than a person who attacks another, thereby creating a conflict between her interests and those of the person attacked. The difference in the degree of responsibility affects the force of the Principle of Fairness-Responsibility: the higher the degree of one's responsibility for a conflict, the stronger the reason that this person should bear its cost.  

49. Brian Barry claims that a similar principle, the "Principle of Responsibility," does not justify substantial economic inequalities, since people are not responsible for many factors that affect differences in the distribution of resources. See Brian Barry, Does Responsibility Undermine Equality? (2003) (unpublished, on file with the Santa Clara Law Review). However, it does apply, inter alia, to isolated situations of the kind described above. Id. at 25 ("It is perfectly fair that those who get themselves into situations in which they need rescuing at great expense and perhaps also risk to life and limb should have to pay for it."). Similarly, Eric Rakowski describes injuries suffered as a result of engaging in dangerous activities as an "exemplary" instance of "option luck" (the consequences of a person accepting an isolated risk that should have been anticipated and could have been declined by that person) that should be born by that person. See RAKOWSKI, supra note 38, at 79.

50. See, e.g., Stephen J. Morse, Diminished Capacity, in ACTION AND VALUE IN CRIMINAL LAW 239, 249 (John Gardner et al. eds., Clarendon Press 1993) (explaining that responsibility is a "matter of degree which ranges along a continuum").

51. This factor is noted, in the context of allocating organs for transplant, by VEATCH, supra note 46, at 314 ("[i]f voluntary, health-risky behaviors are relevant to organ allocation . . . there needs to be a predictable relationship between the amount of the behavior and the degree of risk . . . The degree of risk should be proportional to the amount of the behavior").

52. A different suggestion is that the degree of responsibility should affect
Most people take many risks every day, such as those involved in walking or driving, and these risks might lead to interpersonal conflicts. However, this does not undermine the Principle of Fairness-Responsibility. This principle is significant not only when some individuals are fully responsible and others are in no way responsible for a conflict, but also when there is a difference in the degree of the relevant risk. According to this principle, when several persons are responsible for the existence of a conflict, there is a reason to prefer a more responsible person to another who is responsible to a lesser degree. For example, while every person who drives a car takes a risk of an accident, this risk is relatively minor, and since most people take it, it will work both ways in many interpersonal conflicts that are the result of this risk. In such conflicts, the persons involved would often be responsible to a similar, minor degree. However, driving carelessly or without taking standard precautions, such as wearing a seat belt, is more dangerous and less prevalent, and therefore this risk is more significant. If it would lead to an interpersonal conflict, a careless driver would typically be significantly more responsible than other persons who are involved in the conflict. Driving in motor races is even more dangerous and less common, and thus taking this risk will typically be even more significant.

C. The Proposed Thesis and Self-Defense

Several observations emerge when the proposed set of principles is applied to self-defense as a type of interpersonal conflict, and particularly to the paradigm case of an unjustified and culpable aggressor who endanger the well-being of an innocent victim. In interpersonal conflicts, where a choice has to be made between the well-being of two or more persons, there are several conflicting reasons, expressed in the Well-Being Principle, that should be taken into account, to protect the well-being of each person. When no one is responsible for the existence of the conflict, we should follow the Equal Chance Principle, and thus accord each person the highest possible equal chance to be preferred. This baseline should be rejected only when there is a stronger contrary rea-

the discount rate of the importance of the well-being of responsible persons. See KAGAN, supra note 8, at 55.
son to prefer one of the persons involved.

The Equal Chance Principle needs further elaboration because some commentators have rejected, ignored, or trivialized the idea behind it. Some authors reject the Equal Chance Principle entirely; they consider as immaterial the answer to the question, whom to prefer in interpersonal conflicts in which all pertinent factors are equal or balanced. A similar view is that when choosing between innocent parties, there is no reason to prefer one innocent party to another and therefore it does not matter whom we save, 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." Another view follows the Equal Chance Principle, but does not recognize its significance. It considers the Equal Chance Principle to be only a tiebreaker, i.e., applicable only when all relevant factors are exactly equal. According to both views, any special reason, no matter how trivial, would justify deviating from the Equal Chance Principle.

These views fail to recognize the importance of the reason reflected in the Equal Chance Principle. The choice of whom to prefer in conflicts between innocent parties is not a matter of indifference, even if the considered choice does not "display a vicious preference." Rather, according to the proposed thesis, this choice matters. This is especially true when the competing interests are important, such as in a conflict between lives. When there are two contrary and competing reasons to protect the well-being of different persons, there is important reason, reflected in the Equal Chance Principle, to accord each person the highest possible equal chance to be preferred. The importance of this reason could be vividly seen by imagining oneself in the place of one of the persons whose interest—most notably life—is at stake in an interpersonal conflict. From this perspective, the decision of whom to save is all but inconsequential. This important reason should be outweighed only by a contrary reason whose weight is be-

54. HARRIS, supra note 8, at 71.
55. See id. at 95; John Harris, What is the Good of Health Care?, 10 BIOETHICS 269, 278 (1996).
56. See Broome, supra note 53, at 40; Harris, supra note 55, at 278.
57. HARRIS, supra note 8, at 71.
yond a certain threshold. Thus, the Equal Chance Principle should apply when there is a contrary reason to prefer one of the parties, as long as the contrary reason is not strong enough to cross a certain threshold.

In order to determine the exact nature of this threshold, we need to determine the relative force of the Equal Chance Principle and each competing consideration. Within the proposed framework, there are two considerations which might conflict with the Equal Chance Principle: (1) the reason to prefer a more important interest to a lesser one, which is reflected in the Importance Principle; and (2) the reason to prefer a non-responsible person to a person who is responsible for the existence of the conflict, which is reflected in the Principle of Fairness-Responsibility. The force of each consideration varies in accordance with the feature of each conflict.

The force of the reason reflected in the Equal Chance Principle depends on the importance of the interests at stake: the stronger the competing interests, the stronger the reason to accord each person an equal chance. For example, there is a stronger reason to give an equal chance in a conflict in which lives are at stake than in a conflict between two persons' proprietary interests. The force of the Importance Principle depends on the difference in the importance of the competing interests, taking probabilities into account: the bigger the gap is, stronger is the principle and vice versa. If there is a conflict between the Equal Chance Principle and the Importance Principle, it should thus be resolved in light of the importance of the competing interests in each case. I suggest that, as long as the difference in the importance of the conflicting interests is not substantial, the Equal Chance Principle should prevail over the Importance Principle. However, when the difference in importance increases, the force of the Importance Principle is strengthened and when the difference becomes substantial, the Importance Principle should prevail over the Equal Chance Principle. This conclusion is embodied in the Substantial Difference Principle, which requires that, other things being equal, the person with the substantially more important interest should be preferred. 58

The analysis is more complex when other things are not

58. See generally Segev, Well-Being and Fairness, supra note 9; and Segev, The Concept of Lesser Evil, supra note 9.
equal, and particularly when one or more persons are responsible for the existence of the conflict. The reason to take into account the interests of each person applies also to the interests of persons who are responsible for 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 whose 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 provides an additional reason to prefer the non-responsible person.

The effect of this reason on the analysis of an interpersonal conflict that one person is responsible for its existence, and particularly of the paradigm case of self-defense, in which one person unjustifiably and culpably attacks another who in no way contributed to the conflict, is as follows. When the competing interests are of equal importance, the Well-Being Principle entails two competing reasons of equal importance. In this case, the Equal Chance Principle provides a reason to accord each person the highest possible equal chance to be preferred. However, if the responsible person's degree of responsibility is significant, the reason entailed by the Equal Chance Principle is outweighed by the reason entailed by the Principle of Fairness-Responsibility, to prefer the non-responsible person. The condition of significant responsibility exists by definition in the paradigm case of self-defense. Therefore, the Principle of Fairness-Responsibility provides a basis for the justification to harm the responsible aggressor in order to save the non-responsible person in this case.

The situation becomes even more complicated when the interests of the aggressor and of the non-responsible person are not equally important. In this case, the Importance Principle also factors in. The conclusion is obvious when the choice is between the less important interest of an aggressor and the more important interest of a non-responsible person, because the Importance Principle and the Principle of Fair-
ness-Responsibility both provide reasons to prefer the non-responsible person. The difficulty arises when the aggressor's interest is more important than that of the non-responsible person. For example, such a situation arises when an aggressor threatens the property or bodily integrity of another and the only way to neutralize the threat is by killing the aggressor. In such cases, the Principle of Fairness-Responsibility's preference for the non-responsible person clashes with the Importance Principle's preference for the aggressor due to her more important interest.

Generally, there are three possibilities in this kind of case, depending on the relative force of the relevant principles. First, the fairness-responsibility consideration, that favors the non-responsible person, might outweigh the importance consideration, that favors the responsible person whose interest is more important. This might happen not only when the responsibility is significant and the difference in importance is small, but also when the consideration of fairness-responsibility is very strong, as in the paradigm case of self-defense, even when the difference in the important of the competing is substantial.

Second, the force of both considerations might be roughly the same. In this case, the pertinent considerations on either side balance each other out, and we should follow the Equal Chance Principle.

Finally, the importance consideration that favors the responsible person whose interest is more important, might outweigh the fairness consideration that favors the non-responsible person. This might happen either when the responsibility is insignificant or, when it is, as in the paradigm case of self-defense, if the difference in importance is very big. It is hard to determine the exact threshold, but one point is clear. As noted above, with respect to conflicts between the interests of two non-responsible persons, the Substantial Difference Principle states that the Importance Principle should prevail over the Equal Chance Principle when the difference in the importance of the competing interests is substantial. Therefore, when the person whose interest is more important is responsible for the existence of the conflict, and thus the Principle of Fairness-Responsibility provides an additional, clashing, reason to prefer the non-responsible individual, the difference in importance must therefore be more than sub-
This analysis of the conflict between the Principle of Fairness-Responsibility and the Importance Principle thus entails a principle of proportionality in self-defense situations. According to this proportionality principle, while it is justified to harm an aggressor in order to save a non-responsible person, even if doing so requires causing the aggressor more harm than the harm she will otherwise cause to the non-responsible person, this is only true as long as the difference is not beyond the threshold of proportionality (which requires more than the Substantial Difference Principle).

Beyond this general point, it is difficult to demarcate the exact boundaries of the proportionality principle. The main goal of this article is to account for the foundation of this principle. According to the proposed analysis, this principle is entailed by the interplay between considerations of well-being, derived from the Well-Being Principle, and the various considerations of fairness stemming from the Equal Chance Principle, the Importance Principle and the Principle of Fairness-Responsibility. Generally, it seems to me that the proportionality principle should entail, inter alia, the following conclusions with respect to the paradigm case of self-defense. First, it is unjustified to kill or severely injure an aggressor in order to prevent her from damaging the property of another who is not responsible for the conflict, or from trivially injuring a non-responsible person, such as scratching her. However, second, it is justified to severely injure or even kill an aggressor who would otherwise cause a substantial bodily injury to a non-responsible person. It is also justified, third, to destroy an aggressor's valuable property if she would otherwise destroy a relatively minor property of a non-responsible person. But, fourth, it might be unjustified to destroy an aggressor's house in order to prevent her from causing a relatively small damage to a non-responsible person's property (the answer depends on the degree of responsibility and the exact magnitude of the gap in the importance of the competing interests in each case).

Finally, two clarifications are required with respect to conflicts involving non-responsible persons on both sides, including, in addition to bystanders, also innocent threats and innocent shield of threats. First, with respect to non-
responsible aggressors or persons who constitute threats, a person is not required to stand idle in front of an attack by an innocent aggressor or a person who constitute an innocent threat. Rather, the agent should follow the principles governing the resolution of conflicts between non-responsible parties, particularly the Equal Chance Principle and the Substantial Difference Principle. For example, the agent might be justified in restraining an innocent aggressor or in preventing a threat posed by an innocent person, not only when this could be done without harming the aggressor or the person who constitute the threat, but also in two other cases. First, when this minimizes the harm to all the persons involved in the conflict, in accordance with the first part of the Equal Chance Principle. And, second, when this involves a substantially lesser harm to the innocent aggressor or threat compared to the harm that would otherwise be caused to the other party.

The second clarification is that there might be situations in which an act of defense against an unjustified and responsible aggressor will not only prevent harm to an innocent victim, but also endanger other non-responsible persons (innocent bystanders, innocent threats or innocent shields of threats). In such cases, there is not only a conflict between the aggressor and the non-responsible victim, but also between the non-responsible victim and the other non-responsible persons whom the defensive action might adversely affect. The latter conflict should also be decided in light of the principles governing the resolution of conflicts between non-responsible parties, namely, in such situations, a defensive force that involves harm to non-responsible parties might be justified only if the conditions of the Equal Chance Principle and of the Substantial Difference Principle are met with respect to each of the non-responsible persons.

IV. THE NOTION OF RESPONSIBILITY AS A BASIS FOR SELF-DEFENSE

In its justification of self-defense, the thesis of this article relies heavily on the notion of responsibility. This section further develops this feature of the proposed thesis by evaluating it in light of other theories that justify self-defense based on a notion of responsibility. The first part of this section
presents other responsibility-based views. The second part discusses the main differences between these views and the suggested thesis. In light of this analysis, the final part of this section explores various objections to the notion of responsibility and to the views presented in the first part, in order to determine whether they apply to the proposed theory and, where applicable, whether they are valid.

A. Similar Responsibility-Based Accounts of Self-Defense

Several authors have put forward the idea that responsibility, mainly in the sense of culpability, is a relevant factor in the justification of self-defense. Particularly, Phillip Montague, George (Kai) Draper and Jeff McMahan suggest accounts of self-defense that are similar to the Principle of Fairness-Responsibility in this regard, but differ from it in several important respects.

Montague suggests that in situations in which some individuals will unavoidably be harmed, and it is possible to choose who, it is right to sacrifice the one "whose fault it is" that the situation exists. Montague develops this idea as follows:

when unavoidable harm is 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 paribus) that the harm be distributed

---

60. See discussion infra Part IV.A.
61. See discussion infra Part IV.B.
62. See discussion infra Part IV.C.
63. Beyond the works I discuss in detail below see KAGAN, supra note 8, at 54; Larry Alexander, Self-Defense, Justification, and Excuse, 22 PHIL. & PUB. AFF. 53 (1993); Michael Otsuka, Killing the Innocent in Self-Defense, 23 PHIL. & PUB. AFF. 74 (1994); Coleman & Ripstein, supra note 46, at 94-96. Coleman and Ripstein argue for a “principle of fairness,” according to which “each person should bear the costs of her activities.” They suggest that this principle should be explicated in light of a normative conception that considers the relative importance of the relevant interests, namely, security and the liberty to engage in various activities. They point out that this principle is shared by both libertarianism and liberal egalitarianism but based on different conceptions of responsibility: a causal conception and a choice-based conception, respectively. Coleman and Ripstein further note that libertarianism and liberal egalitarianism subject their different responsibility-based principles to different default rules. According to libertarianism, in the absence of causal responsibility, the costs of misfortune should lie where they fall, whereas according to liberal egalitarianism, misfortunes that are nobody's responsibility are to be held in common.
among those who are blameworthy. 65

Similarly, Draper suggests the following two principles. First, "other things 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." 66 Second,

[other things 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. 67

Finally, McMahan states 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 costs of his wrongful action to be imposed on the morally innocent. 68

Montague, Draper and McMahan subject their principles to similar qualifications. First, the harm to the responsible person must be necessary in order to prevent harm to non-responsible others. 69 This qualification is obvious and uncontroversial as it does no more than ensure that an interpersonal conflict indeed exists.

The second qualification, which is less obvious, is a requirement of proportionality. This requirement - which, as explained in the previous section, is also part of the proposed thesis 70 - is not clearly explained by these authors. Montague suggests

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


67. Id. at 77.

68. McMahan, supra note 8, at 259.

69. Montague, Punishment and Societal Defense, supra note 65, at 33; MONTAGUE, PUNISHMENT AS SOCIETAL DEFENSE, supra note 65, at 46; Draper, supra note 66, at 78; McMahan, supra note 8, at 262.

70. See discussion supra Part III.C.
proportional to the harm that would be suffered by an innocent person under a different distribution. 71

According to Montague, this condition precludes causing major harm to a blameworthy individual in order to prevent a minor harm to an innocent individual. 72

Draper supports his suggestion for a requirement of proportionality with the explanation that "a consideration of welfare overrides the consideration of fairness in this case." 73

According to Draper, "proportionality requires only that the harm inflicted be comparable to (or less than) the harm thereby avoided." 74

Lastly, McMahan asserts that "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." 75

The last qualification suggested by Montague and Draper limits their fault-based distributive principles in order to justify harming only culpable individuals. Montague states that his principle does not "straightforwardly" apply to situations in which harming a blameworthy individual also involves harming an innocent individual. 76 And Draper suggests a "divide-and-conquer strategy," 77 which justifies a single action based on several principles: some of the harmful effects of the action are justified on self-defense grounds, while other harmful effects are justified by other principles. 78

71. MONTAGUE, PUNISHMENT AS SOCIETAL DEFENSE, supra note 65, at 45.
72. Montague, Punishment and Societal Defense, supra note 65, at 33; MONTAGUE, PUNISHMENT AS SOCIETAL DEFENSE, supra note 65, at 46.
73. Draper, Fairness and Self-Defense, supra note 66, at 79.
74. Id.
75. McMahan, supra note 8, at 259 n.11.
76. Montague, Punishment and Societal Defense, supra note 65, at 33; MONTAGUE, PUNISHMENT AS SOCIETAL DEFENSE, supra note 65, at 46.
78. See Draper, supra note 66, at 80-81; Draper, supra note 77, at 66.
Draper notes two principles that might be pertinent in this respect: the principle of double effect and the principle that it is justified to harm innocent persons if this will prevent much more severe harm to other innocent persons. The first principle is incompatible with the theory I suggest. See Segev, The Concept of Lesser Evil, supra note 9. The second principle is similar to the Importance Principle I suggest. However, Draper's main formulation of his principle refers to the number of innocent persons harmed while the Importance Principle does not assume that there is an independent moral significance to the number of individuals involved in interpersonal conflicts. See Segev, The Significance of Numbers, supra note 9. This principle could be applied, more plausibly in my opinion, to conflicts involving just one non-responsible person on each side.
B. Differences between the Proposed Thesis and Similar Responsibility-Based Justifications for Self-Defense

The basic idea underlying the responsibility-based accounts of self-defense mentioned in the previous sub-section is similar to that behind the Principle of Fairness-Responsibility. Accordingly, in several important respects these accounts lead to conclusions that are similar to those of this article with regard to the significance of responsibility in the resolution of interpersonal conflicts. Most obviously, these accounts typically equate innocent aggressors, innocent threats, and innocent shields of threats with innocent bystanders. These accounts also lead to the conclusion 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.

Besides these similarities, there are also significant differences between the views of the three authors described above and this article’s thesis. This subsection explains the most significant differences.

1. The General Framework: The Notion of Responsibility and Other Considerations

This article maintains that a proper resolution of all interpersonal conflicts must rest on an analysis of the interaction of several general, and possibly conflicting, reasons for action. The above three responsibility-based accounts are narrower. They focus almost exclusively on the fault-based principles they suggest. Although these theories mention other, sometimes competing, considerations and qualify these principles in the ways described above, in order to accommodate these considerations, they typically do not explain the

79. See generally Nancy Davis, Abortion and Self-Defense, 13 PHIL. & PUB. AFF. 175, 188-94 (1984); McMahan, supra note 8, at 266-68 (discussing the moral equivalence of various categories of non-responsible persons).

80. Examples are persons who coerce others to harm third parties or who are responsible for a situation in which a choice needs to be made between their interests and those of others at a later time or at a distant place—relative to the time and place in which the responsible party performed the relevant action. See McMahan, supra note 8, at 258; Draper, supra note 66, at 80-81; Draper, supra note 77, at 63-68 (the notion of responsibility includes failure to prevent a conflict if the agent has an “obligation” to do so). See also Lawrence Alexander, Self-Defense and the Killing of Noncombatants: A Reply to Fullinwider, 5 PHIL. & PUB. AFF. 408, 410-12 (1976); Paul H. Robinson, 2 CRIMINAL LAW DEFENSES 73 (1984).
source of these clashes and how they should be resolved.\textsuperscript{81}

Moreover, the positions of Montague and McMahan are unclear with respect to the conflict between their fault-based principles and considerations relating to the overall balance of well-being. On the one hand, as just noted, 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, in various contexts, both seemingly view their fault-based principles as absolute. McMahan states that if his fault-based principle is accepted within a consequential theory, it should have an “absolute priority in the evaluation of outcomes” so that it outweighs “all other features of the possible outcomes.”\textsuperscript{82} Thus, he claims, a defender is justified in killing a culpable aggressor even if there is a “relatively low” probability that the threat the aggressor poses would materialize. Likewise, he argues that a defender is justified in killing “any number” of culpable aggressors if this is necessary to save her life.\textsuperscript{83} Similarly, Montague justifies harming persons who are at fault for an interpersonal conflict, “no matter how many” such persons are harmed as a result, when it is necessary to save the life of a non-responsible person.\textsuperscript{84} (Both McMahan and Montague assume that the number of individuals involved in interpersonal conflicts is in itself a morally significant factor.\textsuperscript{85})

These assertions are incompatible with a principle of proportionality, although both McMahan and Montague formally embrace such a principle.\textsuperscript{86} On the other hand, according to the account I suggest, the Principle of Fairness-Responsibility might clash with other principles, particularly

---

\textsuperscript{81} They avoid this mainly by focusing on one kind of interpersonal conflict: conflicts between equally important interests—primarily lives—of several individuals of which some are responsible for the situation.

\textsuperscript{82} McMahan, supra note 8, at 261.

\textsuperscript{83} Id.

\textsuperscript{84} Montague, supra note 64, at 215; MONTAGUE, PUNISHMENT AS SOCIETAL DEFENSE, supra note 65, at 43.

\textsuperscript{85} See McMahan, supra note 8, at 260-61 (making this claim explicitly in this context); MONTAGUE, PUNISHMENT AS SOCIETAL DEFENSE, supra note 65, at 39 (considering explicitly the numbers as relevant in interpersonal conflicts that are not one’s fault in another place). This assumption is not obvious. See Segev, The Significance of Numbers, supra note 9.

\textsuperscript{86} See discussion supra Part IV.A.
the Equal Chance Principle and the Importance Principle.\textsuperscript{87} And, as explained above, the interaction of these three principles is the basis for a requirement of proportionality.\textsuperscript{88}

2. The Foundation of the Moral Significance of Responsibility: Deontology, Consequentialism and Fairness

Another difference between the thesis of this article and the above mentioned responsibility-based accounts of self-defense concerns the foundation of the idea that, other things being equal, the person responsible for an interpersonal conflict should bear its cost.

The thesis of this article relies on a different foundation, which is generally suggested as an alternative both to deontological and rights-based theories as well as to common forms of consequentialism.\textsuperscript{89} As previously explained, the proposed principles elaborate two basic ideas: (1) the fundamental value and source for reasons for actions in the context of interpersonal conflicts is individual well-being; and (2) the resolution of interpersonal conflicts should reflect several conceptions of fairness relating to the distribution of benefits and burdens: the Equal Chance Principle, the Importance Principle and the Principle of Fairness-Responsibility. This framework is different from both standard deontological theories and from standard consequential theories.\textsuperscript{90}

The above mentioned authors mostly fail to discuss the foundation of their responsibility-based accounts of self-defense. The only brief reference to this issue is McMahan's claim that his account is incompatible with a consequential foundation and "is best interpreted" and in fact "must . . . be understood as a deontological theory."\textsuperscript{91} However, the two grounds noted by McMahan for classifying his principle as deontological are doubtful.

The first ground is that the sacrifice of the culpable party is a just action, but not an action that is right in light of its

\textsuperscript{87} The general explanation of how clashes between the Principle of Fairness-Responsibility and other principles should be resolved is included in Segev, The Concept of Lesser Evil, supra note 9.
\textsuperscript{88} See discussion supra Part III.
\textsuperscript{89} See discussion supra Part I.
\textsuperscript{90} See discussion supra Part III.A.
\textsuperscript{91} McMahan, supra note 8, at 260-63.
consequences, that is, "because it produces a better or more just outcome." But, as McMahan himself immediately notes, it is possible to view responsibility or fault as an intrinsic element in a consequential balance. It is indeed common to refer to considerations of distribution and fairness as non-consequential or "deontological." However, considerations of distribution and fairness, as opposed to deontological constraints on consequential balance, are not necessarily opposed to consequentialism. The conflation of these different considerations seems to be based on the common, but misleading, tendency to encompass all non-utilitarian considerations under the heading of deontology.

McMahan nevertheless rejects the consequentialist interpretation of his idea. He argues that it could "capture our intuitions," or "commonsense morality," only if it accepts justice as having an "absolute priority in the evaluation of outcomes" by outweighing "all other features of the possible outcomes." This, he claims, is because, according to "commonsense morality," a person may defend herself against a culpable aggressor even if this would be detrimental for the balance of the good, such as, for example, killing an aggressor whose lethal threat has only a "relatively low" probability of success. As noted above, however, this view implies a denial of any proportionality requirement and should be rejected, even if the "commonsense morality" endorses it.

The second reason why McMahan considers his idea as deontological reflects the core of standard deontology: he determines which actions are justified, culpable, and excused partly in light of standard deontological constraints on consequential balance. These include the deontological prohibition against intentionally killing an innocent person in order to prevent the intentional killing of several innocent persons.
However, it should be noted that a similar prohibition might be supported by consequentialist considerations, as well as by a non-deontological framework of the kind suggested in this article, if the number of individuals involved in interpersonal conflicts is not in itself a morally significant factor in the adjudication of interpersonal conflicts. 89

3. The Equal Chance Principle versus a Presumption against "Shifting Harms"

The different foundations of the theory suggested in this article, including the Principle of Fairness-Responsibility, and McMahan's fault-based principle lead to a difference in the background principles, namely, the principles that apply when no one is responsible for an interpersonal conflict. The background principle is important also to cases in which there is a person who is responsible for the existence of the conflict, as in the paradigm case of self-defense, since, in such cases, the responsibility-based principle clashes with the background principle.

The background principle I suggest is the Equal Chance Principle. On the other hand, McMahan suggests that the background principle 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 to happen. 100 This distinction is incompatible with the non-deontological basis of this article's thesis and particularly with the Equal Chance Principle. Moreover, as I explain elsewhere, I believe that other possible grounds for a presumption against shifting harms should be rejected as well and consequently I endorse a wide interpretation of the Equal Chance Principle that applies as the background principle in all interpersonal conflicts. 101

---

89. This issue is discussed in Segev, The Significance of Numbers, supra note 9.

100. McMahan, supra note 8, at 252-53. Similarly, Otsuka argues that if a danger to one person can be prevented only by harming another it is unjustified to act. See Otsuka, supra note 63, at 76-77.

101. This point is explained further in Segev, The Concept of Lesser Evil, supra note 9. This is another demonstration of the nature of the thesis proposed in this article, which rejects deontological distinctions but replaces them with the Equal Chance Principle that is incompatible with standard forms of conse-
This difference is especially significant since the proposed thesis considers the Equal Chance Principle as reflecting an important reason for action rather than as a mere tie-breaker. Therefore, only strong contrary reasons can outweigh this reason. For example, Draper argues that when two people need an object in order to survive and one of them is the owner of the object, the latter should be preferred. On the other hand, according to the thesis I suggest, ownership itself may not provide a reason that is strong enough to outweigh the Equal Chance Principle, especially when this principle is strong, as is the case in a conflict involving important interests, such as life.

4. The Scope of the Principle of Fairness-Responsibility

Another important difference between the Principle of Fairness-Responsibility and the principle suggested by Montague and McMahan is their scope. The Principle of Fairness-Responsibility relies on a choice-based conception of responsibility, and therefore applies to both wrongful actions and justified actions that lead to interpersonal conflicts. On the other hand, both Montague and McMahan limit their principles to wrongful actions. Montague refers to "fault" and to "culpable" actions and defines these terms to include "intentional, reckless or negligent" actions, namely, unjustified, and perhaps unexcused, actions that lead to interpersonal conflicts. Similarly, McMahan explicitly limits his principle to morally "culpable" parties, which also connotes unjustified and unexcused actions. However, he encounters difficulties in trying to support this view. He explores several explana-
tions for the common assumption that there is a different, supplementary justification, more limited in nature, for preferring innocent victims over innocent aggressors and innocent threats, but not over innocent bystanders.\textsuperscript{107} However, he concludes that he is “unable to find a fully convincing justification” for this assumption.\textsuperscript{108} He suggests that the reason for this assumption might be the tendency to “overgeneraliz[e] 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, McMahan assumes, reflects this assumption.\textsuperscript{109}

Draper’s position in this respect seems closer to the one suggested in this article. Although he initially refers to “moral fault” or “blame,” he later expands his analysis 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.”\textsuperscript{110} However, Draper seemingly intends this formula to apply more broadly than the suggested Principle of Fairness-Responsibility because he adopts a bias towards the status quo by considering an effort to avoid a danger, for which the agent is not responsible, as a benefit in the relevant sense. Thus, Draper illustrates his principle with the following example: a person who is justifiably acting in self-defense strikes an undercover police officer who is trying to hit him. According to Draper, this person benefits in the relevant sense by his action although he acts justifiably.\textsuperscript{111} The Principle of Fairness-Responsibility, on the other hand, would view both parties in this case as equally non-responsible, provided

\begin{thebibliography}{99}
\bibitem{107} Id. at 263-90.
\bibitem{108} Id. at 288.
\bibitem{109} Id. at 288-90.
\bibitem{110} Draper, \textit{supra} note 66, at 84.
\bibitem{111} Id. at 84. Draper also claims that individuals who create risks to others should compensate them, if the risks materialize, even if creating the risks was justified and not blameworthy. \textit{See also} Devlin, \textit{supra} note 103, at 246-47. \textit{Compare} Otsuka, \textit{supra} note 63 at 90-91 (arguing that it is justified to harm a person who is responsible for a threat to another even if the latter is free of blame, for example, if the threatening party mistakenly thinks that the other is about to harm him). According to the thesis I suggest, it is unjustified to harm persons who are innocent threats if they did not take any special risk that led to the conflict.
that the police officer is not responsible for the existence of the initial danger.\footnote{112} This is not to say that the police officer is justified in striking the other non-responsible person, but rather that the proper course of action is to follow the Equal Chance Principle unless the expected harm to one of the parties is substantially less serious than the expected harm to the other, in which case the Substantial Difference Principle requires that we prefer the latter.

\section{5. Agent-Neutral versus Agent-Relative Principles}

Additional differences stem from disagreements that are more general in nature. The principles suggested in this article aim to provide general, agent-neutral, guidance to all agents in the resolution of all kinds of interpersonal conflicts. On the other hand, both Montague and McMahan make several distinctions in this respect between different agents.

Montague claims that in standard self-defense situations, one is “not simply permitted to kill in self-defense, but one has a right to do so,” and that a third party is “obliged to choose in favor of the person whose life is threatened.”\footnote{113} He argues further that 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.”\footnote{114} McMahan similarly claims that self-defense in the strict sense (namely, an agent who defends herself) should be “optional,” because “one is not always required” to protect herself against a culpable aggressor.\footnote{115} These controversies reflect complex issues that are beyond the scope of this article.\footnote{116} Therefore, I will only note two points here. First, the proposed principles are all agent-neutral, and apply equally to all agents in all kinds of interpersonal conflicts. Second,
according to the proposed principles, no substantive line should be drawn between permissions, obligations, rights and duties.

Another possible difference, which also raises general questions that are beyond the scope of this article, is that the Principle of Fairness-Responsibility applies to both benefits and burdens, while the other three fault-based principles refer only to the allocation of burdens. The suggested thesis does not endorse the distinction between burdens and benefits, or between changing the position of individuals for the worse or for the better. In my opinion, these distinctions are not morally insignificant in themselves, although there might be practical considerations for adopting similar distinctions in certain contexts.

C. Criticism of the Significance of Responsibility in the Justification of Self-Defense

Despite the above differences between the Principle of Fairness-Responsibility and the other responsibility-based principles discussed, the basic idea underlying all these principles is similar. It is therefore instructive to consider the criticisms of the other responsibility-based principles and their relevance to the thesis of this article.

1. Fairness versus Retribution

Several critics mistakenly reject responsibility as a ground for resolving interpersonal conflicts due to its supposed "retributive character." As pointed out above, the justification to prefer non-responsible over responsible persons in interpersonal conflicts is not necessarily retributive. In

117. See discussion supra Part IV.A.
119. See discussion supra Part IV.A.
120. See discussion supra Part IV.B.
121. See David Wasserman, Justifying Self-Defense, 16 PHIL. & PUB. AFF. 356, 372 (1987); UNIACKE, supra note 4, at 185-86; GEORGE P. FLETCHER, BASIC CONCEPTS OF CRIMINAL LAW 145 (1998) ("[L]inking 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 contrasts punishment with self-defense.
122. See discussion supra Part III.B.
fact, none of the fault-based principles mentioned above is retributive. Specifically, the Principle of Fairness-Responsibility represents a distributive consideration, based on the concept of fairness, which reflect what is considered the right way of allocating an unavoidable burden, and not on a retributive notion.

2. The Validity of the Principle of Fairness-Responsibility

David Wasserman criticizes Montague's responsibility principle by claiming that greater fault is neither a sufficient nor a necessary condition for choosing one life over another. First, Wasserman asserts that greater fault is insufficient to justify self-defense: he argues that it is not always justified to sacrifice the person who is most at fault in a forced-choice situation. He gives several examples to demonstrate this claim. One example involves a choice between two people who will die without a transplant of a liver: where the first is not responsible for his predicament, but is old and expected to live only six months even with the new liver, and the second is young and otherwise healthy and is expected to live thirty years with the new liver, but is responsible for his need of a new liver due to his excessive alcohol consumption. In another of Wasserman's examples, one must choose which of the passengers of a drowning ship should be saved, after one of them ran the ship into an iceberg, intending to kill another man who had raped and caused the death of his daughter. Wasserman avoids claiming that, in these examples, it is clearly wrong to sacrifice the responsible persons (the young alcoholic and the rapist) rather than the non-responsible persons, but he does argue that "[e]ven if we are inclined to favor the sacrifice of the culpable parties, it is not clearly the 'dominant' solution as it is in standard self-defense" situations.

Wasserman second claim is that greater fault is also un-

---

123. See Montague, supra note 105, at 88; McMahan, supra note 8, at 259-60; Draper, supra note 77, at 66.
124. See discussion supra Part III.B.
125. Wasserman, supra note 121, at 366-69.
126. Id. at 366.
127. Id. at 367.
128. Id.
129. Id. at 368.
necessary to justify self-defense, as it might be justified to sacrifice an aggressor who is not the most responsible person for the attack.\textsuperscript{130} 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 follow that we should help the aggressor if the only options are his death or the death of the provocateur.\textsuperscript{131}

Wasserman is right that greater fault, in the sense of responsibility for the existence of an interpersonal conflict, is neither a sufficient nor a necessary condition for sacrificing individuals. However, this view seems compatible with Montague's thesis, and, in any case, it is certainly compatible with the thesis of this article and particularly with the Principle of Fairness-Responsibility. This Principle does not always justify the sacrifice of the person who is most responsible for an interpersonal conflict, but rather holds that there is always a reason that the person responsible for an interpersonal conflict should bear its cost instead of a non-responsible person. This reason is not necessarily decisive within the framework suggested in this article. On the contrary, since it might conflict with other reasons for action, incorporated in the other proposed principles, the resolution of interpersonal conflicts in which this reason is applicable depends on its force relative to the force of the other applicable conflicting considerations.\textsuperscript{132}

This may be demonstrated in the examples given by Wasserman himself. Although there is indeed room for hesitation with respect to the resolution of all three examples, this does not undermine the thesis suggested in this article and particularly the Principle of Fairness-Responsibility. On the contrary, in all of these examples there is a reason that the persons responsible for the situation will be sacrificed rather than each of the other, non-responsible, persons. However, in all three examples, this reason is weaker than in the paradigm case of self-defense,\textsuperscript{133} and since it sometimes

\textsuperscript{130}  \textit{Id.} at 366.
\textsuperscript{131}  Wasserman, \textit{supra} note 121, at 368.
\textsuperscript{132}  \textit{See} discussion \textit{supra} Part III.
\textsuperscript{133}  With respect to the ship example, Michael Gorr asserts that the "justification for killing those whom Wasserman terms 'past aggressors' is every bit as strong as our justification for killing present aggressors (all else equal)."  \textit{See}
clashes with other reasons, it is not necessarily decisive in these cases. Particularly, the following points should be taken into account with respect to the relative force of the competing reasons in these examples.

First, in all three examples, the responsibility of the "culpable parties" for the conflict is much weaker than that of a typical aggressor. This is most obvious in the ship example where the father is described as "[o]vercome with rage."\textsuperscript{134} This is also true in the transplant example, in which it is reasonable to assume that the young person did not cause his illness and thus the conflict deliberately, but only negligently or, at most, recklessly. As for the third example, provocation is a classic ground for considering persons less responsible for aggression and, consequently, less culpable (although provocation is typically considered as a basis for mitigation of punishment, its effect on responsibility is also relevant to the notion of fairness that relies on responsibility).\textsuperscript{135} Thus, in all three examples, the reason to sacrifice the responsible person is significantly weaker 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.\textsuperscript{136}

Second, in the provocation example there is room for doubt the validity of the comparative judgment that a person who provokes another, even by "relentless insults and harassment,"\textsuperscript{137} is more culpable than a person who reacts to such a provocation in a way that forces a choice between their lives. This doubt is especially strong with respect to culpability in the relevant respect, that is, responsibility for the existence of the conflict of lives. In fact, it seems that the contrary might well be true. The provocateur's character might well be profoundly flawed, and he is certainly responsible for

Gorr, supra note 6, at 265. This proposition is too strong, as explained in the text.

\textsuperscript{134} Wasserman, supra note 121, at 367.

\textsuperscript{135} But see generally JEREMY HORDER, PROVOCATION AND RESPONSIBILITY (1992) (providing a critical analysis of the relation between provocation and responsibility).

\textsuperscript{136} With respect to the ship example, see Montague, supra note 105, at 82-83. Moreover, as pointed out by Gorr, in real life versions of the examples in which the responsibility 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, supra note 6, at 265.

\textsuperscript{137} Wasserman, supra note 121, at 368.
the conflict to a significant extent. However, it seems that a person who faces mere insults, and no physical danger, but nevertheless chooses to react with deadly force bears more responsibility for the existence of the situation in which a choice must be made between his life or that of the person who insulted him.\(^{138}\)

Finally, in the transplant example, this article's thesis suggests that - in addition to the reason entailed by the Principle of Fairness-Responsibility, to prefer the young, responsible person over the old, non-responsible person - there is also a contrary reason, based on the Importance Principle, to prefer the young person who will live with the new liver for thirty additional years, to the old person, who will live with it only six more months. This is so since it is reasonable to elaborate the Importance Principle so that, other things being equal, the interest of a person in living for thirty additional years would be considered as more important than the interest of a person in living for only an additional six months. Admittedly, this interpretation of the Importance Principle is controversial because it raises complex questions about the value of life.\(^{139}\) However, it does seem a plausible interpretation. And, if this interpretation is indeed correct, it supports saving the younger person on the ground that his interest in getting the transplant is stronger.\(^{140}\) If this interpretation is indeed correct, there are, in the above example, two, clashing considerations: the reason to prefer the non-responsible older person and the reason to prefer the person who will benefit most from the transplant. It is hard to determine the relative force of these competing considerations. But the presence of the contrary reason to prefer the young person explains why at least some people feel a dilemma in the transplant example. If there is indeed such a contrary reason, it affects the overall judgment concerning the example without casting doubt on the validity of the Principle of Fairness-Responsibility, that is, on the idea that there is a reason to prefer the old as the non-responsible party.

---

138. See Montague, supra note 105, at 86.

139. This issue is discussed in Segev, Well-Being and Fairness in the Distribution of Scarce Health Resources, supra note 9.

140. 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. See Montague, supra note 105, at 85.
3. The Principle of Fairness-Responsibility and Deontological Constraints

Suzanne Uniacke raises a related, but distinct, criticism against the notion of responsibility as a basis for resolving interpersonal conflict:

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 grounds that I can 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 a transplant.141

Uniacke acknowledges, however, "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."142 Uniacke does not explicitly explain the difference between the two cases, but it appears that her view is that the idea that responsibility should affect the resolution of interpersonal conflicts should be limited by deontological, agent-relative constraints, which forbid positively harming persons regardless of other considerations, most notable consequential considerations but perhaps also other, non-consequential, considerations.

This criticism of the idea that responsibility should affect the resolution of interpersonal conflicts is, however, misguided.

The first point that it is important to emphasize is that even if we assume that there are valid deontological considerations, such as the distinction between positively harming persons and saving persons from harm, this does not undermine the Principle of Fairness-Responsibility, for several reasons. First, it should be remembered that the Principle of Fairness-Responsibility states that there is a reason to prefer a non-responsible person to a responsible one, rather than that this reason is necessarily and therefore always decisive. Therefore, the validity of principle could not be undermined by the claim that there are also other valid considerations, in-

141. UNIACKE, supra note 4, at 188-89 (internal citations omitted).
142. Id.
cluding deontological consideration, such as the distinction between positively harming and not helping. Generally, the existence of a certain consideration does not annul other considerations, but rather implies that there are several pertinent considerations in certain cases.

This observation is often overlooked due to the tendency to describe deontological considerations as “absolute,” a tendency which is reinforced by common expressions such as “there are things that one must not do regardless of the consequences” or “rights are trumps.” However, even if deontological considerations are decisive, this does not imply that other considerations are not valid. It could only imply that other considerations would be outweighed by deontological considerations whenever both kinds of considerations are present and therefore these other considerations would not affect the practical decision in these cases.

Moreover, it is also important to bear in mind that even if deontological considerations are valid they are not necessarily decisive. Indeed, it seems that even those who support deontological considerations do not typically consider them as absolute.

The second point that should be emphasized is that, even if deontological considerations are not only valid but also “absolute,” namely, outweigh all other considerations in all the cases in which they apply, the Principle of Fairness-Responsibility can not only be relevant but can even be decisive in other cases, in which deontological consideration do not apply, like in Uniacke’s example of the two endangered persons that only one of whom can be saved and only one of whom is responsible for the danger threatening the other.

Finally, it should be noted that it is not at all clear that there are valid deontological constrains such as those based on the distinction between positively harming and not helping. To be sure, there is a natural repulsion from harming persons positively, especially when this involves removal of their vital organs and thereby killing them, as in the heart variation of Uniacke’s transplant example. This repulsion explains why it is easier to accept the Principle of Fairness-Responsibility when the question is which endangered person to save when not all can be saved, rather than whether to

143. See RONALD DWORKIN, TAKING RIGHTS SERIOUSLY (1977).
positively kill one person in order to save another. However, it is difficult to identify a cogent normative basis for distinguishing the two situations.\textsuperscript{144} Indeed, even some of those who believe that there is a morally significant difference between killing and letting die acknowledge that it is difficult to rationally defend this distinction.\textsuperscript{145}

The assumption underlining the thesis suggested in this article is that there are no valid deontological considerations and particularly that the difference between positively harming and not helping is not in itself a morally insignificant one.\textsuperscript{146} Accordingly, I believe that the Principle of Fairness-Responsibility is universal, namely, apply to all interpersonal conflicts. According to the proposed thesis, if a person is responsible for endangering the life of another, and the only way to save the latter involves sacrificing the former by positively harming him, there is a reason to do so and, if there are no other pertinent considerations, or if there are only weaker applicable considerations, it is justified to positively sacrifice a responsible person in order to save a non-responsible person.\textsuperscript{147}

4. Former versus Present Responsibility

After arguing against responsibility as both a sufficient and a necessary condition for preferring one individual to another in the resolution of interpersonal conflicts, Wasserman’s suggests his own responsibility-based view of the justification of self-defense. Wasserman’s view is based on what he considers the 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 present aggressor whose responsibility is exhibited at the time the conflict is resolved. He argues that it

\textsuperscript{144} Various grounds for distinguishing the two kinds of cases are examined and rejected in Segev, \textit{The Concept of Lesser Evil, supra} note 9.

\textsuperscript{145} \textit{See Jeff McMahan, Killing, Letting Die, and Withdrawing Aid}, 103 \textit{ETHICS} 250, 279 (1993) (arguing that although the considerations underlying the deontological distinction between killing and letting die are not cogent, the intuitive force of this distinction is so strong that it is impossible to imagine morality without them).

\textsuperscript{146} \textit{See discussion supra} Part III.A.

\textsuperscript{147} Note that, according to the proposed thesis, it is the responsibility for the danger, rather than the fact that the responsible person is a “criminal,” which is the relevant consideration.
might be justified to harm the latter in self-defense, but not the former.\textsuperscript{148}

However, Wasserman fails to explain persuasively why this distinction is morally significant.\textsuperscript{149} While one's responsibility for the existence of an interpersonal conflict seems to be of obvious moral significance (although it is not necessarily decisive), the time when such responsibility accrues lacks apparent normative significance. Time may be important only as an indication of responsibility. When responsibility is exhibited in present aggression, i.e., at the time the conflict is resolved, the agent, who has to resolve the conflict, typically has relatively clear evidence of the identity of the responsible person and the nature of her responsibility. On the other hand, typically the agent will have less evidence of prior responsibility.\textsuperscript{150} This evidentiary difference explains the hesitation to harm "prospective aggressors" pointed out by Wasserman, while casting doubt on his argument for the independent moral significance of present aggression.\textsuperscript{151}

In support of his distinction between present and former responsibility, Wasserman also briefly 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."\textsuperscript{152} However, this assertion, which is not supported by further argument, is also unpersuasive.

5. Justification versus Evading Responsibility

Seumas Miller also criticizes Montague's suggestion that fault is relevant in the adjudication of interpersonal conflicts.\textsuperscript{153} 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."\textsuperscript{154} However, there is no indication that Montague had this idea in mind when he formulated

\textsuperscript{148} Wasserman, supra note 121, at 371-77.

\textsuperscript{149} See Gorr, supra note 6, at 267 ("[I]n 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 discussed in the next section, but rejects the arguments offered for it by Wasserman. See UNIACKE, supra note 4, at 185.

\textsuperscript{150} See Gorr, supra note 6, at 265.

\textsuperscript{151} Wasserman, supra note 121, at 373-74.

\textsuperscript{152} Id. at 375-76.

\textsuperscript{153} Miller, supra note 114, at 240-41.

\textsuperscript{154} Id.
his suggestion. In any case, this is not the idea underlying this article's thesis: 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 her action. Instead, the idea is that, a rational agent should consider the aggressor's responsibility as a morally significant factor in the resolution of the conflict. The significance of responsibility is not in that it relieves the agent from responsibility but rather in that it provides a reason to prefer the interests of a non-responsible person to the interests of a responsible person.\textsuperscript{155}

A similar argument, from a different perspective, is made by Cheyney Ryan. He argues that no justification for self-defense is needed, because while the attacker is responsible for a "situation in which someone's life will be lost," the victim "is not responsible for this situation, it is merely presented to him."\textsuperscript{156} This argument is misguided for the same reason. Harming an aggressor is an action that requires justification, whether it is performed by the person who is being attacked or by a third person. Although the aggressor forces a choice, this does not relieve the person who harms him from responsibility for his action. Therefore, a justification for this harm is required. As suggested above, it could and should be justified with reference to the aggressor's responsibility for the existence of the conflict.\textsuperscript{157} But there is a difference between evading responsibility for harm and justifying harm. Justification of harm assumes, rather than denies, responsibility for it.

6. \textit{Justification versus Description}

Miller criticizes Montague's notion of responsibility, as a basis for self-defense, by arguing 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, and, as a mere description, it lacks

\textsuperscript{155} See discussion supra Part III.
\textsuperscript{156} Cheyney C. Ryan, \textit{Self-Defense, Pacifism, and the Possibility of Killing}, 93 ETHICS 508, 515-16 (1983). Ryan also raises the possibility that the mere fact that an attacker \textit{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 \textit{positive} grounds could be given for this principle." See id. at 516-17.
\textsuperscript{157} See discussion supra Part III.C.
normative force. But this criticism too is unfounded. The notion of responsibility, as reflected in Montague's theory and in the Principle of Fairness-Responsibility is not an empirical description of a situation but rather a normative idea, which gives preference to the person who is not responsible for a conflict over a person who is. This principle might require additional argument, but it is obviously not merely a descriptive one.

7. Responsibility and Intention

Finally, Miller criticizes Montague's suggestion on the ground that, in referring only to the fact that the aggressor forces a choice on the agent between the aggressor himself and the defender, it ignores a crucial factor: the aggressor's intention to attack, and harm, another. These factors, he continues, are distinct since a choice could be forced unknowingly and both should be taken into account. The gist of this criticism is unclear. If it attempts to distinguish between a causal and a choice-based conception of responsibility, it is, in my opinion, valid. However, such an interpretation does not undermine Montague's view, which specifically refers to a culpability-based conception, rather than to a causal conception, of responsibility. Similarly, this interpretation does not undermine the thesis suggested in this article, which adopts a choice-based conception of responsibility and, accordingly, distinguishes between responsible and non-responsible persons even though the latter could too be "forcing a choice" in some (causal) sense.

Moreover, the Principle of Fairness-Responsibility reflects degrees of responsibility in evaluating the force of the reason it incorporates. The force of the reason to prefer the non-responsible, or the less responsible, person varies in accordance with the degree of responsibility of the other responsible or more responsible person. Therefore, according to this conception, causing a conflict intentionally involves a greater degree of responsibility than doing so, for example, accidentally or even negligently.

158. Miller, supra note 114, at 240-41.
159. Id. at 241.
160. See discussion supra Part III.B.
V. NON-RESPONSIBILITY-BASED ACCOUNTS OF SELF-DEFENSE

This section compares the proposed justification for self-defense to several other dominant justifications of self-defense, which, unlike the theories discussed in the previous section, do not rely on the notion of (choice-based) responsibility, and then explains why the proposed justification is superior to these views.

I will not discuss all possible accounts of self-defense. Particularly, I will not explore several other accounts of self-defense that relate to the basic assumptions of the proposed theory. I put aside agent-relative accounts of self-defense that are based on the idea that agents are permitted to prefer their interests to those of others. Second, I ignore the deontological claim that self-defense is justified since it does not involve intending harm (or evil) in light of the principle of double effect. Third, I do not discuss utilitarian accounts of self-defense. And, finally, I mostly disregard contractual accounts of self-defense. As noted above, and explained more fully elsewhere, the proposed thesis is agent-neutral, rejects deontological distinctions, is not utilitarian and does not rest on contractual ideas.

161. For agent-relative conceptions of self-defense, see Jeremy Waldron, Self-Defense: Agent-Neutral and Agent-Relative Accounts, 88 CAL. L. REV. 711 (2000). The possibility of an agent-relative justification for self-defense is also mentioned by Nancy Davis with respect to conflicts between innocent parties. She argues that since there is no moral difference between the persons involved in such conflicts, the only basis for preferring one person over the other is an agent-relative permission for each to prefer his interests (although the well-being of each is equally important and it is not fairer to prefer one to the other). See Davis, supra note 79, at 188-94.

162. For an analysis of the applicability of the principle of double effect to self-defense, see KAGAN, supra note 22, at 132-44; UNIACKE, supra note 4, at 92-155.

163. For a brief discussion of the nature of such accounts, see KAGAN, NORMATIVE ETHICS, supra note 8, at 251-52. For a relativist contractual account of self-defense see Gilbert Harman, Moral Relativism Defended, 84 PHIL. REV. 3, 9, 14-15 (1975) (morality is an unconscious and implicit agreement between a group of people to adhere to certain principles and the rights to self-defense and self-preservation derive their relative moral force from the fact that people cannot rationally form the intention not to defend themselves against aggression). See also discussion infra Part V.A.3.

164. See discussion supra Part III.

165. Segev, The Concept of Lesser Evil, supra note 9. For a rejection of agent-relative justifications for self-defense on other grounds, see McMahan, supra note 8, at 268-71.
Another preliminary point that should be noted concerns accounts of self-defense that rely on the concept of "lesser evil." In its most rudimentary sense, the concept of lesser evil—namely, the idea that it is sometimes justified to positively harm or to sacrifice the interest of one person in order to prevent harm to another—is merely an empty framework, which lack any substantive content and could accommodate any normative theory, including the thesis suggested in this article and the other theories discussed in it. Therefore, accounts of self-defense that rely on the notion of lesser evil in this general form are not helpful. However, the concept of lesser evil is often understood more narrowly to reflect a specific moral theory, most commonly, a utilitarian conception. This interpretation is misguided to the extent that it implies that utilitarianism is the only possible foundation for the rudimentary concept of lesser evil.

The self-defense theories explored at length in this section include rights-based accounts, accounts that rely on a comparison of the value of the interests of aggressors relative to the value of the interests of the persons they attack, and accounts that are based on the causal responsibility, notably the notion of intrusion. These theories provide various interpretations of the rudimentary concept of lesser evil in the context of self-defense, that are different from the one suggested in this article.

A. Theories of Rights

Rights-based accounts of self-defense typically assume that persons have rights to life, bodily integrity, property and other interests, which include, at least, the right that others will not positively harm these interests. Self-defense then

166. See, e.g., GEORGE P. FLETCHER, RETHINKING CRIMINAL LAW 857-60 (Little, Brown & Co. 1978); ROBINSON, supra note 80, at 69-72; Omichinski, supra note 7, at 1452-53 (all mentioning the concept of lesser evil as a possible rationale for self-defense).

167. For associations of the concept of lesser evil with utilitarianism, see generally FLETCHER, supra note 166, at 790-91; Omichinski, supra note 7, at 1456, 1458.

168. See discussion infra Part V.A.

169. See discussion infra Part V.B.

170. See discussion infra Part V.C.

171. See Bedau, supra note 7, at 559 ("[T]he right of self-defense, according to all Natural Rights theorists, including Blackstone, is simply a special case of the right to life."). For Blackstone's understanding of the right to life, as one
JUSTIFICATION FOR SELF-DEFENSE

is justified, according to these accounts, as a protection of the relevant rights from attack.\(^{172}\)

Reliance on the notion of rights involves three main preliminary problems, in my opinion.\(^ {173}\) First, the notion of rights blurs the basic question of the justification of actions—whether actions properly reflect all relevant reasons for actions—by focusing instead on issues such as whether actions violate duties or rights. This is so since the concepts of duties and rights are significant when they are understood as primal concepts, which precede and affect the issue of justification of actions, as opposed to merely represent conclusions concerning justification in certain terminology. Second, I believe that the notion of rights should be rejected if understood as a fundamental moral concept that deviates from the idea that individual well-being is the fundamental value. I suggest that the value of individual well-being should be supplemented or adjudicated in light of considerations of fairness, but not replaced by other, competing, fundamental values, such as duties and rights. Finally, and perhaps most importantly in the context of self-defense, rights-based accounts often fail to explain the foundations of the alleged rights. Particularly, rights-based accounts of self-defense typically assume that self-defense is justified rather than explain why this is the case.\(^ {174}\)

\(^{172}\) component of “the right of personal security,” which includes also the protection of “limbs,” “body,” “health,” and “reputation,” see 1 WILLIAM BLACKSTONE, Commentaries *129.

\(^{173}\) I ignore the problem of identifying the relevant right. For this problem, see Kaufman, supra note 5, at 23:

The dilemma for the rights theory is that any single purported right which is broad enough to cover all or most of the situations in which defensive force is justified will be too open-ended to be helpful. Yet simply to assert a list of distinct rights to justify distinct uses of defensive force is no more helpful, as it provides no unifying explanatory account, but only an ad hoc list.

\(^{174}\) See Segev, The Concept of Lesser Evil, supra note 9 (for a further explanation of these points).

\(^{174}\) 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, 37 PHIL. STUD. 405-17 (1980) (“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”); Kaufman, supra note 5, at 20 (“whatever might be the correct account of self-defense, rights theory will not do—it is at best unhelpful, and at worst positively pernicious as an explanation of the permissibility of defensive force”; id. at 26 “One cannot avoid a deep suspicion of circularity. What is the direction of explana-
Rights-based accounts of self-defense may be classified in accordance with their answer to a basic question that faces every theory aimed at justifying self-defense based on the notion of rights, namely, why protecting the rights of individuals by thwarting attacks is not forbidden as a violation of the attackers’ equivalent rights. The first answer to this question is that while individuals generally have a right not to be harmed, attackers lack this right. Variations of this answer differ in the explanation they offer as to why and when, namely, with respect to which attackers, this is the case. The second answer is that, like people in general, attackers have a right not to be harmed. However, this right clashes with the equivalent right of the people they attack. According to this account, this clash should be resolved in favor of the latter, since their rights are more important and thus override those of the attackers. Variations of this answer again differ in the explanation they offer as to why and when this is
The third answer to the above question differs in nature and rests on the idea that the right to self-defense is against the state. Each of these answers is considered in turn.

1. Do Attackers Lack the Right Not To Be Harmed?

a. Justification for Harm and Rights

One explanation why attackers lack the right not to be harmed is that persons have only the right not to be harmed unjustifiably, and that harming attackers in self-defense does not violate this right as harm in self-defense is justified. This view is sometimes called the “specification” view of rights. The problem is that this account most obviously begs the question of justification: it clearly assumes that self-defense is justified rather than explain why this is so.

b. Forfeiture and Loss of Rights

A second variation of the view that aggressors do not have a right to life, bodily integrity, liberty or other pertinent interests, and therefore do not have a right not to be harmed with respect to these interests, is based on the idea that persons can forfeit these rights by their own conduct. Accord-

---

176. See discussion infra Part V.A.2.
177. See discussion infra Part V.A.3.
178. See Judith Jarvis Thomson, A Defense of Abortion, 1 PHIL. & PUB. AFF. 47, 57 (1971); Parent, supra note 174, at 405-17 (labeling this account the “moral specification” view of rights).
179. See Thomson, Self-Defense and Rights, supra note 174, at 37-39; Wasserman, supra note 121, at 362; Ryan, supra note 156, at 513. Indeed, even William Parent—who defends this view, although only by pointing out flaws in other accounts of rights—acknowledges that (like any theory of rights, in his opinion) it does not offer a justification of actions. See Parent, supra note 174, at 417.
180. The idea that rights, including the right to life, can be forfeited by a wrongful or criminal act (though not voluntarily alienated) is mentioned by John Locke and by William Blackstone in various contexts. Locke mentions it, for example, when he writes, while explicating the institution of slavery, that a person could “by his fault, forfeit his own Life, by some act that deserves Death.” John Locke, An Essay Concerning Civil Government in TWO TREATISES OF GOVERNMENT 284 (Cambridge Univ. Press 1988) (1690). Blackstone similarly writes that life may be “frequently forfeited for the breach of those laws of society, which are enforced by the sanction of capital punishment.” See generally 1 WILLIAM BLACKSTONE, COMMENTARIES *133; 4 WILLIAM BLACKSTONE, COMMENTARIES *178. For analysis of Blackstone's view, regarding forfeiture of rights and its applicability to self-defense, see Bedau, supra note 7, at 567-68.
ing to a narrow version of this view, attackers forfeit their rights to the relevant interests, and thus their right to be free of harm, by unjustifiably, and culpably, attacking and endangering others.\textsuperscript{181} A broader version of this view is that individuals forfeit or lose their right to life, and thus their right not to be killed, if they pose "an unjust immediate threat to another person's life or proportionate interest."\textsuperscript{182} This view is broader since it includes not only culpable aggressors but also blameless aggressors and even people who do not act at all but merely pose involuntary or passive threats, "provided the threat sufficiently resembles an assault, a clear case being Nozick's man thrown down the well."\textsuperscript{183}

The broad version of the view that attackers lack the right to be free of harm is offered by Judith Jarvis Thomson and elaborated by Uniacke. Both authors begin with the assertion that, other things being equal, every person has a right to life, which entails, 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. However, they claim that there is an exception to this rule: persons who attack others or constitute threats (according to Uniacke: "unjust immediate threats") to others, whether culpably or innocently, violate the right of others not to be killed. Therefore, such attackers lose their right not to be killed. Consequently, the authors conclude, it is justified to kill aggressors and people who constitute threats, regardless of whether they are responsible for attack or threat under a choice-based conception of responsibility, in order to protect the persons who are being attacked

\textsuperscript{181} See Jeffrie G. Murphy, The Killing of the Innocent, 57 The Monist 527, 547 (1973) ("those attacking have by the attack forfeited certain of their rights"); Joel Feinberg, Voluntary Euthanasia and the Inalienable Right to Life, 7 Phil. & Pub. Aff. 93, 103-04, 111-12 (1978) (the forfeiture idea should be contrasted with the idea that killing innocent threats, such as "innocent shields," is justified but involves an infringement, although not a violation, of their right to life); A. J. Ashworth, Self-Defence and the Right to Life, 34 Cambridge L.J. 282, 288-89, 294 (1975) (making both the claim that attackers' rights are forfeited and the claim that they are overridden, without distinguishing between them); Kagan, supra note 8, at 173 ("we might say that although each of us has a right not to be harmed, if I unjustifiably attack another I forfeit my right"); Omichinski, supra note 7, at 1461 (the forfeiture of rights account of self-defense applies only to culpable aggressors and not to innocent attackers).

\textsuperscript{182} Uniacke, supra note 4, at 196.

\textsuperscript{183} Id. at 229.
or endangered by them. This exception, they explain, does not generally apply to bystanders who are not related to the attack or threat, although harming them might be justified based on other exceptions to the rule against harming and particularly killing people. Thomson and Uniacke both consider this view as essentially equivalent to the forfeiture theory.

c. The Problem with Forfeiture and Loss Accounts

The accounts of self-defense that rely on the idea that attackers lose, forfeit, or otherwise lack the right not to be harmed, could be based on two background assumptions. First, these accounts could provide a complete justification for self-defense if the assumption is that it is permissible to perform any action that does not violate rights of persons. Alternatively, and more plausibly, these accounts can provide merely as a partial justification for self-defense, based on the view 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.” Under this view, the idea of loss or forfeiture of rights is nevertheless part of the justification of self-defense since it removes one moral constraint on harming individuals—their rights—and thus explains why self-defense does not wrong aggressors and is therefore an exception to the general prohibition on harming persons. In other words, “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.” This account therefore does not purport to justify self-defense in itself. It merely suggests

184. Judith Jarvis Thomson, Self-Defense, 20 PHIL. & PUB. AFF. 283 (1991); Uniacke, supra note 4, at 156, 209-27 (noting the similarity of her view and Thomson’s). In other places, Thomson suggests that it is also justified to kill innocent shields of threats. See Thomson, Self-Defense and Rights, supra note 174, at 38; Thomson, supra note 10, 370-71; but see Uniacke, supra note 4, at 160-72.

185. Uniacke considers this variation “as theoretically on a par” with a theory of forfeiture of rights. Uniacke, supra note 4, at 195. In her words, “there need be no substantive difference” between them. Id. at 208-09. Thomson writes 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. Thomson, Self-Defense, supra note 184, at 301.

186. Uniacke, supra note 4, at 191.

187. Id. at 191-92, 203-04.

188. Draper, supra note 66, at 88.
that if there is a different, independent justification to harm attackers in self-defense, we should not be concerned about their rights.

On the other hand, the first account, that does purport to provide a complete justification for self-defense, raises an obvious concern. This account seem to imply that, since culpable as well as innocent aggressors and innocent threats lose their rights not to be harmed, their well-being or interests are of no value and deserve no consideration. For if having certain rights is a necessary (albeit not necessarily a sufficient) condition to the moral status of individuals, then those who lack the required rights are deprived of any moral status. This entailment seems misguided; even if aggressors deserve less consideration in the resolution of the interpersonal conflicts they are responsible for, it is implausible to regard them as lacking any moral status.

In particular, if the prohibition against harming people relies at least in part on the right of individual not to be harmed, and if aggressors and persons who constitute threats lack this right, there seems to be no reason why such aggressors and threats would not be harmed gratuitously and without limits. Therefore, the view that aggressors and threats lack the right not to be harmed appears incompatible with the idea that there must be good reasons for harming persons and with the requirements that only necessary and proportional force should be used in self-defense. After all, if aggressors and persons who constitute threats lack rights, and if rights are at least part of the basis for the prohibition on harming persons, it seems that there is no bar to harming them without good reason and to using unnecessary and disproportionate force against them.

Thus, commentators have criticized the forfeiture theory on the ground that it unwarrantably finds it justified for agents who are unaware of the attack, and thus act without a

189. A more radical, but unexplained, criticism is that it “is inconsistent for society to acknowledge the right to life as supremely important and then render a wrongdoer’s life as worthless.” See Omichinski, supra note 7, at 1466.

190. See Bedau, supra note 7, at 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”); Kaufman, supra note 5, at 23. This is indeed the implication of the assertion, made by Murphy, that since attackers forfeit their rights “there is no conflict” in “interfering with them.” See Murphy, supra note 181, at 547.
good reason, to harm aggressors. Similarly, critics also fault the forfeiture theory because it condones those who act beyond the extent needed to repel the attack. It was also argued that the same point applies to the view that there is a deontological constraint against doing harm, which does not apply to aggressors.

As explained above, the idea behind this criticism applies more widely. First, it undermines any account of self-defense that is based on the claim that attackers lack rights that prevent other persons from harming them. Therefore, this criticism applies not only against the forfeiture theory but also against the specification variation of the view that attackers do not have the right not to be harmed. Second, this criticism applies not only to the requirement to use only necessary force, but also the requirement to use only proportional force. If aggressors lack the 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 (i.e., the interest of the person attacked to avoid the harm the aggressor might cause her and the interest of the aggressor to avoid the harm necessary to prevent him for causing this harm)?

191. See Fletcher, supra note 3, at 144-45.

192. For example, after the attack is over or when it could be stopped by less harmful, and perhaps non-harmful, means (this point was noted also by Thomson herself in a previous article). See Sanford Kadish, Respect for Life and Regard for Rights in the Criminal Law, 64 CAL. L. REV. 871, 884 (1976); Thomson, Self-Defense and Rights, supra note 174, at 34-37; Parent, supra note 174, at 412-13, 415-16; Joshua Dressler, Rethinking Heat of Passion: A Defense in Search of a Rationale, 73 J. CRIM. L. & CRIMINOLOGY 421, 454 (1982); Wasserman, supra note 121, at 361; Ryan, supra note 156, at 511-12; Draper, supra note 66, at 87-88; Omichinski, supra note 7, at 1456, 1466; Kaufman, supra note 5, at 24, 25-26. 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 this view permits the employment of unnecessary force. See KAGAN, supra note 22, at 136-37.

193. See Kagan, supra note 8, at 92-93 (“if the constraint against doing harm simply didn’t apply in cases where the person being harmed is a deliberate aggressor, then neither of these requirements [that limit self-defense to the use of necessary and proportional force], would make any sense”).

194. Both Thomson and Uniacke endorse the proportional force requirement, even with respect to culpable aggressors. See Thomson, Self-Defense and Rights, supra note 174, at 285-86, 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, supra note 4, at 143-55, 213, 229.
One answer is that aggressors lack the right not to be harmed only to the extent necessary to stop their attack and prevent the danger it poses.\textsuperscript{196} In other words, according to this reply, aggressors and persons who constitute threats retain the right not to be harmed when harming them is unnecessary.\textsuperscript{196} This argument may be expanded to accommodate also a requirement of proportionality, which is applicable at the time an aggressor constitutes a risk. It might be claimed that aggressors lose or forfeit their right not to be harmed only to the extent that is needed for proportional protection, and retain the right not to be harmed in a non-proportional way.\textsuperscript{197}

The problem with this answer is that the limitations it places on the implications of the loss or forfeiture of the rights of aggressors seem ad hoc, since there is no explanation as to why the loss or forfeiture of the relevant rights is revoked along the lines they suggest (so that only necessary and proportional force is justified in self-defense).

A variation of the above answer is sometimes applied to the right to life. According to this view, aggressors lose or forfeit important rights only when their aggression threatens an important interest. For example, William Blackstone writes that aggressors forfeit their right to life only when they are expected to use \textit{lethal} force or perform another crime punishable by death. According to this view, when aggressors are expected to perform only lesser harms (and, consequently, offenses), they retain their right to life (although they might forfeit lesser rights) and therefore could not be justifiably killed in self-defense.\textsuperscript{198} Blackstone suggests this explanation

\begin{footnotes}
\item[195] This answer is suggested by Ashworth and Uniacke. Ashworth, \textit{supra} note 181, at 289; UNIACKE, \textit{supra} note 4, at 213.
\item[196] Ashworth, \textit{supra} note 181, at 289; UNIACKE, \textit{supra} note 4, at 213.
\item[197] This indeed seems to be the unarticulated assumption behind the views of Ashworth and Uniacke. \textit{See} Ashworth, \textit{supra} note 181, at 296 (stating that the principle of proportionality is entailed by the attackers' right to life); \textit{id}. at 303 ("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"); UNIACKE, \textit{supra} note 4, at 196 (discussing the view that individuals forfeit their right to life if they pose "an unjust immediate threat to another person's life or proportionate interest").
\item[198] \textit{See} 1 WILLIAM BLACKSTONE, COMMENTARIES *133 (stating that life may be "frequently forfeited for the breach of those laws of society, which are enforced by the sanction of capital punishment"); Omichinski, \textit{supra} note 7, at 1454-55 (citing Blackstone as an example for this position).
\end{footnotes}
with reference to the law of eighteenth-century England, but one may interpret his view to state a broader principle that refers to the penalties determined by each legal system's prevailing law. It is also plausible to interpret this view as assuming that the relevant offenses are not only punishable by death, but also should be treated this way, if this is ever the case. Indeed, arguably, the only relevant condition is the latter.

Blackstone's view hints to a possible explanation with regard to limitations placed on the forfeiture of rights: aggressors who perform an offense punishable by death lose their right to life according to the law. However, this explanation only raises the same substantive question in a different formulation: is the relevant law justified?

This line of response thus highlights a more general problem shared by other accounts of self-defense that rely on the notion of rights: these accounts fail to justify their conclusion because they do not provide an independent explanation of the basis of their construal of the rights they invoke. For this reason, it seems that these accounts beg the important questions and, in fact, are merely ways of expressing, rather than justifying, normative conclusions.

\[ d. \text{ Overall Justification and Justice with Respect to Rights} \]

Uniacke acknowledges the circularity of rights-based accounts of self-defense. She attempts to resolve this problem by distinguishing between two normative categories: what is

199. 1 WILLIAM BLACKSTONE, COMMENTARIES *133; 4 WILLIAM BLACKSTONE, COMMENTARIES *178.

200. See McMahan, supra note 8, at 277-78 (stating that Thomson's view "is an ingenious exercise in begging the question" and that the above response entails a strange view that determines the scope of rights in light of contingent factors such as the options of defense); Draper, supra note 66, at 86-89 (stating that Thomson does not provide a justification for her view). Compare George P. Fletcher, Defensive Force as an Act of Rescue, 7 SOC. PHIL. & POLY 170, 178 (1990). As 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. There are also others problems with accounts of self-defense that rely on the idea of forfeiture. For example, their treatment of the case of putative self-defense: it is hard to see why an innocent person, who is rationally but mistakenly considered to constitute a threat, loses her right not to be harmed. See Kaufman, supra note 5, at 24. I explain why such cases of putative self-defense should be considered justified in Segev, supra note 5.
justified (and accordingly unjustified) overall and what is just (and accordingly unjust)—a notion that she interprets to refer only to violations (or non-violations) of rights. In light of this distinction, she defines the scope of the right to life only in terms of what is just as opposed to what is justified. Under her view, actions can be just but unjustified, for example, when an agent uses necessary and proportional force, including lethal force, against an unjust aggressor but is either unaware of the relevant circumstances or acts with a bad motive. Such actions, she claims, do not violate the aggressor’s right not to be harmed, but are nevertheless unjustified due to the mental state of the agent. Similarly, she argues, actions can be unjust if they involve a violation of rights but justified, if the violation is the only way to prevent a serious harm to others. 201

This distinction between the notions justified and just can avoid circularity, and provide a basis for Uniacke’s assumptions regarding the scope of self-defense, only if an independent account of the notion of just (and unjust), and, accordingly, of the content and scope of rights, is provided. However, despite of Uniacke’s extensive discussion of “the self-defense justification of homicide,” 202 and notwithstanding the crucial role the distinction between justice and justification plays in her theory concerning the basis for lethal self-defense, she writes that such an account “is something that I cannot realistically attempt here” and that this account “remains important, unfinished business.” Uniacke’s omission is of special significance since she rejects two kinds of “indirect” accounts that, in her opinion, “have notable contemporary advocates” and “prima facie both seem plausible candidates for generating the required exception”: “a contract theory” and “a two-level theory of optimality,” essentially, rule-utilitarianism. 203 In support of her account of the justification of self-defense, Uniacke thus eventually relies only upon what she thinks is the “considerable intuitive plausibility” of her assumptions concerning the scope of self-defense. 204 Therefore, Uniacke suggests only an unrealized possibility of a non-circular theory that supports her conclusions, and, in fact, de-

201. UNIACKE, supra note 4, at 210-18, 230-31.
202. This is the subtitle of her book, id.
203. Both accounts are suggested briefly by KAGAN, supra note 22, at 135.
204. UNIACKE, supra note 4, at 218-28.
rives these conclusions in a circular way from intuitions regarding the scope of self-defense.\(^{205}\)

Furthermore, Uniacke's basic distinction and some of her assumptions about the scope of self-defense seem perplexing in themselves. First, her claim that there are two fundamental normative dimensions—what is "justified" overall and what is "just" in terms of rights—rather than one substantive normative category that takes account of all the relevant considerations—seems to reflect an unexplained assumption that there is something unique about rights in comparison to other normative factors. More specifically, it is difficult to find a convincing reason for considering the causal responsibility of innocent aggressors and innocent threats as morally significant. The claim made by Thomson and Uniacke is that even an innocent, immediate attack or threat violates the right of the person in danger not to be harmed. However, this view is unexplained and seems unconvincing since, by definition, the harm that innocent aggressors and innocent threats are expected to cause is not due to any choice they made.\(^{206}\)

### 2. Overriding the Rights of Attackers?

Thomson mentions, in addition to her forfeiture theory, also another version of a rights-based justification for self-defense. According to this version, attackers retain the right not to be harmed, even during the attack, but this right is overridden by the rights of the people they endanger. Specifically, she suggests that an aggressor has a right not to be killed but this right is not absolute and is overridden by the

---

205. See Kaufman, supra note 5, at 21 (Uniacke's account "serves as a restatement rather than an explanation or justification of self-defense").

206. Compare this criticism with Thomson's view that there is a morally significant difference between, on the one hand, innocent aggressors, innocent threats and innocent shields of threats, and, on the other hand, innocent bystanders. See Alexander, supra note 63, at 53-66; Otsuka, supra note 63, at 74-94 (arguing that innocent aggressors or threats do not violate rights); McMahan, supra note 8, at 266-68, 275-76 (stating that Thomson's position, that justified attacks involve violation of rights, entails the odd conclusion that justified attackers lose their rights not to be harmed and to defend themselves); Parent, supra note 174, at 405-17; Draper, supra note 66, at 86-90. Compare Jerome Hall, General Principles of Criminal Law 436 n.85 (2d ed., Bobbs-Merrill 1960) (claiming that protection against an insane attacker is akin not to defense against a culpable aggressor but rather to protection from "a natural force," since "an insane person is not bound by duties of the penal law" and therefore "he cannot violate any legal right conferred by that law").
victim's equivalent right not to be killed. However, Thomson herself seems to acknowledge that this idea does not provide a justification for self-defense, but only assert that self-defense is justified, as long as it is not accompanied by an explanation of why, and consequently when, attackers' rights are overridden.

3. A Right against the State?

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

207. Thomson, Self-Defense and Rights, supra note 174, at 42-47. Compare Bedau, supra note 7, at 569 (claiming that "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 act[s]" of the right bearer, "scarcity of resources" and the need to resolve conflicts of rights).

208. See Kadish, supra note 192, at 884-88.

209. Id.

210. Id. at 885. Waldron interprets Kadish's view as agent-relative. See Waldron, supra note 161, at 747.
threats, in self-defense. Therefore, this account only states that if there is such a right, it does not disappear after the establishment of a state. However, even if it is unjustified for a state to prevent persons from defending themselves or others by harming culpable or innocent aggressors, this does not explain why such harm is justified.\footnote{Kadish's account has been advocated as the most plausible account of self-defense on the ground that it “can generate rational rules of criminal conduct,” such as limiting self-defense to necessary and proportionate force, as the right to resist aggression is derived from the state, and therefore “the state logically can define its scope narrowly to preserve important societal goals.”\footnote{But see UNIACKE, supra note 4, at 222.} But although “the state logically can define” the scope of the right of self-defense in many ways, in order to promote various goals, Kadish's account does not explain how this right should be molded.}

Kadish’s account has been advocated as the most plausible account of self-defense on the ground that it “can generate rational rules of criminal conduct,” such as limiting self-defense to necessary and proportionate force, as the right to resist aggression is derived from the state, and therefore “the state logically can define its scope narrowly to preserve important societal goals.”\footnote{But see UNIACKE, supra note 4, at 222.} But although “the state logically can define” the scope of the right of self-defense in many ways, in order to promote various goals, Kadish’s account does not explain how this right should be molded.

B. Value-Based Accounts of Self-Defense

1. The Value of the Interests of Aggressors

A different kind of justification for preferring one individual to another in interpersonal conflicts is based on a comparison of the value of their interests. Shelly Kagan points out specifically that one way to capture the view that non-responsible persons should be preferred to those who are responsible for the existence of a conflict is by discounting the importance of the well-being of responsible persons, the discount rate being determined in accordance with the degree of responsibility.\footnote{Omichinski, supra note 7, at 1467.} In this spirit, several commentators argue that self-defense is justified based on a comparison of the relative value of the interests of attackers and those of the people they attack. According to this view, self-defense is justified since the interests of the former are less valuable than those of the latter.

John Harris claims that, in contexts involving conflicts of lives, self-defense against an unjustified and culpable aggressor is justified because by attacking another person the aggressor “is giving the world a reason to value his life less than

\footnote{KAGAN, supra note 8, at 55.}
that of others." This, he explains, is because the aggression itself "represents an attack on the very value that must support any defense of the value of [the aggressor's] own life." Therefore, "[a]n aggressor could not ... consistently claim that his own life should be respected while attempting to kill his victim."214 Harris 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."215

George Fletcher provides a related account of self-defense, which he attributes mainly to Anglo-American law. This account, he explains, 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.216

These accounts lead to two conclusions regarding the scope of self-defense. First, since these accounts seem to apply only to culpable aggressors, they do not justify the use of defensive force against innocent aggressors and innocent threats.217 Second, with respect to culpable aggressors, these accounts seem to entail a requirement of proportionality, which is indeed required by Anglo-American law,218 because the interests of culpable aggressors, although discounted, still should supposedly receive some consideration.

214. HARRIS, supra note 8, at 69.
215. Id. at 77.
216. FLETCHER, supra note 166, at 857-58; George P. Fletcher, Proportionality and the Psychotic Aggressor: A Vignette of Comparative Criminal Theory, 8 ISR. L. REV. 367, 377 (1973); Fletcher, supra note 3, at 142; GEORGE P. FLETCHER, A CRIME OF SELF-DEFENSE: BERNARD GOETZ AND THE LAW ON TRIAL 24 (1988). See also Dressler, supra note 192, at 454:

[S]elf-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 which flows from his death is negated or, at least, reduced.

It is not clear whether Dressler refers only to culpable aggressors or also to innocent attackers and innocent threats.

217. HARRIS, supra note 8, at 71-73.
218. See Ashworth, supra note 181, at 296-97; ROBINSON, supra note 80, at 81-88; PAUL H. ROBINSON, 1 CRIMINAL LAW DEFENSES 86-88 (1984).
2. The Problem with Value-Based Accounts of Self-Defense

The problem with value-based justifications of self-defense is that they do not explain why, and to what extent, an aggression diminishes the value of the aggressor's interests. The theory proposed in this article suggests that the responsibility of an aggressor for the conflict is a reason to sacrifice her interests rather than those of another, but this is not because her interests are of no value, or are less valuable, than those of other people. The interests of all persons, including culpable aggressors, have value, and therefore there is a reason not to harm even culpable aggressors. However, the responsibility of aggressors for conflicts they cause, between their interests and the interests of the persons they attack, is the basis for a competing reason that the responsible aggressors, rather than their victims, should bear the inevitable cost of the conflict. This argument presents, in my opinion, a more straightforward and plausible explanation as to why self-defense is justified than the argument that the interests of aggressors are less valuable than those of other people.\(^{219}\)

Harris' reference to the inconsistency in the hypothetical position of aggressors, regarding the importance of each party's well-being,\(^ {220}\) even if correct, does not in itself support the conclusion that it is justified to harm aggressors in self-defense. First, the assumption that the inconsistency in the hypothetical position of culpable aggressors shows that there must be a justification for self-defense is doubtful, since the source of this inconsistency is obvious: aggressors could not consistently both justify their aggression and reject self-defense against them since their aggression is unjustified. Moreover, even if the inconsistency in the hypothetical position of culpable aggressors would have showed that there must be a justification for self-defense, it does not explain the nature of this justification.

Harris' other remark, that aggressors "in a sense . . . volunteered" for any injury that preventing their attack may involve, is even more puzzling.\(^ {221}\) It is hard to see in what rele-

---

219. See discussion supra Part III.C.
220. HARRIS, supra note 8, at 69.
221. Id. at 77.
vant sense this is true, apart from the straightforward normative judgment - not based on any idea of actual consent (as opposed to the normative idea of hypothetical consent) - that it is right that the interests of aggressors, rather than those of non-responsible victims, would be sacrificed when some sacrifice is unavoidable.

3. **Further Criticism of Value-Based Accounts of Self-Defense**

The suggestion that self-defense is justified due to the diminished value of the aggressors' interests is often criticized by claiming that it conflicts with several widely accepted views. Several aspects of this criticism are reasonable, but other aspects imply unwarranted conclusions, in my opinion.

One common accusation is that the diminished-value accounts conflict with the widely held view (which is assumed to be reflected in the law) that resolving 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 a morally better or worse person. For example, according to this view, when it is possible to save only one of two people, who would otherwise both drown, the fact that one of them is a murderer should not be taken into account in deciding who to save.\textsuperscript{222}

This criticism might be correct if interpreted narrowly to apply only to the assessment of the importance of the lives of those involved in the conflict. But the above claim might reflect a wider view, which excludes considerations relating to prior actions of the persons involved in interpersonal conflicts. According to one possible variation of this view, the only relevant consideration in the resolution of interpersonal conflicts is the relative importance of competing interests (assuming that everyone would agree that this is a valid consideration, for example, that in a conflict between one person's life and another's property, the former should prevail). This interpretation should, I believe, be rejected, since it ignores the normative significance of responsibility for the existence of interpersonal conflicts, as reflected in the Principle of Fair-

\textsuperscript{222} See Kadish, supra note 192, at 880-82; Wasserman, supra note 121, at 358-59; Rosen, supra note 7, at 47.
ness-Responsibility.

The argument that self-defense is justified since the interests of aggressors are less valuable than those of others has also been criticized on another ground. The diminished-value argument could account for the common view that it is justified to harm aggressors much more severely than the harm they are expected to cause only if the value of the interests of aggressors would be diminished to an extremely low degree. This seems implausible.

While this criticism seems to raise a valid point, it is unwarranted to the extent that it implies a rejection of a requirement of proportionality in such conflicts. According to the theory suggested in this paper, the expected harm involved in any act, including harm to aggressors, is a reason against the act (this view is expressed in 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 (this point is expressed in the Importance Principle).

Thus, while the Principle of Fairness-Responsibility provides a reason to prefer a non-responsible person over a responsible aggressor, this reason is not necessarily decisive, since it may conflict with a reason based on the Importance Principle. Despite the responsibility of an aggressor for the existence of an interpersonal conflict, it is unjustified, for example, to kill her in order to prevent trivial damage to the property of a non-responsible person or a low probability danger to a minor interest of a non-responsible person.

C. Intrusion-Based Accounts of Self-Defense

1. Unjustified Intrusions

Another common rationale for self-defense is based on the idea that self-defense is justified since it stops an unjusti-

---

223. 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 a morally significant factor. See Segev, The Significance of Numbers, supra note 9. But obviously, this is the assumption underlying this criticism.

224. See Kadish, supra note 192, at 882; Wasserman, supra note 121, at 359; McMahan, supra note 8, at 262.

225. See discussion supra Part III.A.

226. See discussion supra Part III.A.

227. See discussion supra Part III.C.
fied (wrongful) intrusion of aggressors on others. This argument is typically applied to all aggressors, regardless of whether they are responsible for their intrusion under a choice-based conception of responsibility. 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 her, regardless of whether the person who poses the threat is responsible for this attempt.228 This claim resembles Uniacke's suggestion, within her rights-based theory, that self-defense is justified against persons who pose "unjust immediate threats," whether culpable or not.229

A variation of this theme is presented by Fletcher, who associates it mainly with the view of Immanuel Kant and with German and Soviet law. This variation is based on the idea that unjustified attacks "intrude" or "encroach" upon the "living space," "personal domain," or "bodily integrity" of individuals, and thus violate "individual autonomy" and, consequently, the "Legal Order." Fletcher argues that this account supports unlimited self-defense on the ground that "Right need never yield to Wrong." Particularly, he interprets this theory to imply that self-defense is equally justified against any unjustified (wrongful) aggressor, whether culpable or not, and without qualification by a requirement of proportionality. In his words, "killing an aggressor is permissible if it is the only means available to prevent the invasion of even a minor interest."230

Fletcher sometimes qualifies this assertion by distinguishing between the notion of "Right," which completely ignores the interests of wrongful aggressors, and the notion of

---

229. UNIACKE, supra note 4, at 172-93.
230. Fletcher, supra note 3, at 141; FLETCHER, supra note 121, at 136, 144; FLETCHER, supra note 166, at 860-74; Fletcher, supra note 200, at 171-72; FLETCHER, A CRIME OF SELF-DEFENSE, supra note 216, at 32-33, 35-36; Fletcher, Proportionality and the Psychotic Aggressor, supra note 216, at 378-87; George P. Fletcher, Punishment and Self-Defense, 8 LAW & PHIL. 201, 210, 213-15 (1989); George P. Fletcher, The Nature of Justification, in ACTION AND VALUE IN CRIMINAL LAW 175, 181 (John Gardner et al.: eds., Clarendon Press 1993).
"morality" or "justice," which takes account of the interests of aggressors and thus supports a requirement of proportionality. He suggests that although the former notion dominates the law of self-defense, the latter, which he associates mainly with the institution of punishment, affect it as well. Fletcher also notes that both German and Soviet law, despite their alleged reliance on the above account of self-defense, in fact qualify the position that the protection of rights is unlimited by introducing, like Anglo-American law, a requirement of proportionality, for example, through the doctrine of "abuse of rights."232

Charles Fried advocates a similar theory. According to his theory, an "unjustified attack creates a relation . . . of wrong, between the attacker and his victim," which justifies harming the attacker in self-defense. He applies this analysis to all 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 uncontrollably.233

Similarly, Frances Kamm argues that, where a conflict exists between a person who poses 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," she further argues, "is to stand in an inappropriate relation to the other person (even if the imposition is not the result of anyone's unjust act)."234 Kamm leaves open the question of whether her thesis should apply to 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 a "minor injury"

232. FLETCHER, supra note 121, at 136-37, 144-45; Fletcher, supra note 166, at 871-74; Fletcher, Proportionality and the Psychotic Aggressor, supra note 216, at 368, 390; Fletcher, Punishment and Self-Defense, supra note 230, at 181. With respect to German law, see also Mordechai Kremnitzer, Proportionality and the Psychotic Aggressor: Another View, 18 ISR. L. REV. 178, 210-11 (1983).
233. FRIED, supra note 10, at 44-53.
234. Frances Myrna Kamm, The Insanity Defense, Innocent Threats, and Limited Alternatives, 6 CRIM. JUST. ETHICS 61, 65 (1987). Kamm suggests that the person who constitutes the innocent threat is required to stop the threat and perhaps to compensate the other for the harm caused by the threat.
compared to death to the other innocent person. But she does not explain the grounds for this very significant qualification. If this is the entire scope of her theory, then it is very limited indeed. Conversely, the thesis suggested in this article provides that in Kamm’s last example, when the conflict is between life of one innocent person and a “minor injury” to another innocent person, the “minor injury” is justified, in light of the Substantial Difference Principle, regardless of who constitutes the (innocent) threat.

Paul Robinson offers a different variation of the intrusion-based account of self-defense. He suggests that the “balancing of harms” should include not only “physical harms,” but also “intangible evils,” like “the compelling societal interest” in preserving the rights of “bodily autonomy” and “property ownership” of individuals and like the general interest in “condemning unjustified aggression.” This account is reminiscent of Thomson’s view that self-defense is justified because both culpable and innocent aggressions violate the rights of the individuals they attack. Here, however, the emphasis shifts from analysis of rights to the “intangible harms” of entering the personal sphere of people which often accompany physical aggression.

Finally, Michael Gorr does not offer a justification for self-defense, because, as noted in the introduction, he doubts whether this is possible. But Gorr does suggest some principles for the resolution of various conflicts, one of which considers the fact that a person is posing a threat, including innocent threat, to another as a morally significant factor. Gorr attempts to explain the basis for the intuitive distinction between an innocent bystander, whom he thinks may not be justifiably harmed in order to prevent a similar harm to another, and an innocent threat, whom 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

235. Id.
236. See discussion supra Part III.C.
237. ROBINSON, supra note 80, at 70.
238. Thomson, supra note 184, at 283-310.
such measures." Gorr acknowledges that "it is difficult to see why the mere fact that a person plays a causal role in threatening harm to another has any moral relevance whatever" when this person is not responsible for playing such a role. He even concedes 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 was] in some way at fault in bringing about such a threat." Indeed, he concludes by admitting that the question of 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 account for the intuitive distinction between innocent bystanders and innocent threats.\(^2\)

2. The Normative Insignificance of Intrusion

The main problem with all intrusion-based accounts of self-defense is that they fail to explain why an unjustified intrusion on a person justifies harming the intruder, if the latter is not responsible for the situation under a choice-based conception of responsibility. The resolution of interpersonal conflicts should take account of the well-being of all persons involved. When resolving conflicts between non-responsible persons, the only valid reason to prefer one over the other is the relative importance of the competing interests. This reason is strong enough to outweigh the Equal Chance Principle only when there is a substantial gap in the importance of the competing interests.\(^2\) Otherwise, since the well-being of both non-responsible parties is equally important, and fairness does not provide a reason to prefer either, we should follow the Equal Chance Principle. This is, of course, if there is no way of avoiding the conflict, for example, by restraining the aggressor without harming either party.

An "imposition" or "intrusion" on persons is no doubt undesirable, since it typically has a negative impact on the well-being of those who suffer the intrusion. Therefore, there is a

\(^2\) Gorr, supra note 6, at 248, 250, 252-53.
\(^2\) See discussion supra Part III.C.
reason, reflected in the Well-Being Principle, to stop intrusions. But when this involves harming the intruder, there is an additional, contrary, reason, which also stems from the Well-Being Principle that applies also to non-responsible intruders, not to harm the innocent intruder. None of the above accounts explain why the cost of an unjustified intrusion should fall on an innocent intruder who is not responsible for the intrusion. A non-responsible intrusion does not constitute a cogent reason for preferring one innocent person to another. It seems that the views discussed in this section consider only one side of the equation—the reason to protect an innocent party from an attack. But there is also a contrary reason to take account of the interests of attackers, especially when they are innocent too.\footnote{Fletcher's view has been criticized on the similar ground that rights of individuals are not “absolute” but rather “context-sensitive,” and that the resolution of interpersonal conflicts should also take account of other considerations, beyond the reason to prevent intrusions, such as the interests or “rights” of others and the innocence or culpability of others with respect to the conflict at hand. See Kremnitzer, supra note 232, at 183-89; Larry Alexander, Justification and Innocent Aggressors, 33 WAYNE L. REV. 1177, at 1182 (1987). Fletcher's view has also been criticized, in a similar spirit, based on a contractual reasoning and the idea of equality, according to which the interests and rights of aggressors should be taken into account as well. See David A. J. Richards, Human Rights and the Moral Foundation of the Substantive Criminal Law, 13 GA. L. REV. 1395, 1435-36 (1979); David A. J. Richards, Rights, Resistance, and the Demands of Self-Respect, 32 EMORY L.J. 405, 426 (1983) (“[T]he 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 bystanders. See FRIED, supra note 10, at 53; FLETCHER, supra note 166, at 862-64; FLETCHER, supra note 121, at 143-45. The view that any intrusion justifies harming the intruder has also been criticized on the ground that it leads to the odd conclusion that unjustified and responsible aggressors are justified in defending themselves against their victims who fight back. See McMahan, supra note 8, at 257-58 (criticizing what he calls the “Orthodox View,” according to which it is justified to harm anyone who is harming another); Russell Christopher, Self-Defense and Objectivity: A Reply to Judith Jarvis Thomson, 1 BUFF. CRIM. L. REV. 537, 538 (1998) (“[Thomson’s] objective approach, properly considered, not only justifies the threats posed by morally innocent, but causally harmful, actors, but also the force used by culpable (evil) aggressors and unwittingly forfeits the life of morally and causally innocent victims.”).}
still be required.\textsuperscript{242} Indeed, the above reason is the basis for the requirement of proportionality with respect to responsible aggressors, despite the reason to prefer non-responsible persons over them.

As Gorr emphasizes, the distinction between an innocent bystander and an innocent threat seems to have an intuitive appeal. The proper place of intuitions in normative thinking is complex and controversial. Doubtless, reliance on intuition in philosophical and legal discussions is common and often extensive.\textsuperscript{243} But there also seems to be a wide agreement that intuitions are fallible and require rational support.\textsuperscript{244} This article rests on the assumption that intuitions should be justified on rational grounds. If we assume that there are correct normative standards, then normative conclusions should be ultimately based on a rational analysis of reasons rather than on mere intuitions, as there is no reason to assume that intuitions constitute correct normative standards or that they are always guided to them by an invisible hand. There may be some room for intuitions in normative thinking, as a starting point to normative inquiry, but the emphasis, I think, should be on rational reflection.

Notably, intuitive reasoning can yield conflicting conclusions. Specifically, the intuitive judgment regarding the relevant question is not one-sided. Beside the intuition that distinguishes between innocent bystanders and innocent threats, there seems to be also an intuitive support to the more general idea that innocent persons should be treated

\textsuperscript{242} Contrary to Fletcher’s interpretation of the “autonomy” view. See discussion supra Part V.C.1.

\textsuperscript{243} For dominant general examples, see Thomson, \textit{Afterward}, supra note 174, at 257-60; THOMSON, \textit{supra} note 10, at 4-33; KAMM, \textit{supra} note 10; F. M. KAMM, MORALITY, MORTALITY, VOL. 2: RIGHTS, DUTIES AND STATUS (1996).

equally in the resolution of interpersonal conflicts, inter alia, in the sense that one innocent person should not be preferred over another. Therefore a conflict exists in this context between two strong intuitions, and the overall conclusion in the intuitive sphere is unclear. When this is the case, there is all the more reason to turn to rational reflection.

Furthermore, the intuition that a person may defend herself against an innocent attacker, even when this involves inflicting more harm upon the innocent attacker than is acceptable to inflict upon an innocent bystander, could be explained on two alternative grounds that do not rely on the doubtful assumption that intrusion per se is morally significant.

First, intuitions regarding conflicts involving aggression are probably influenced by the fact that in real-life cases most aggressors are culpable and therefore, when the agent does not have information concerning this issue, it is reasonable to assume so based on experience. This rational assumption of culpability justifies preferring the person being attacked over the aggressor in most conflicts, and this conclusion holds true even when the rational assumption turns out to be mistaken in a specific case.245

Second, the intuition that it is permissible to harm innocent threats in defense, more than innocent bystanders in interpersonal conflicts, could at least be partly attributed to the observation that the employment of defensive force should often be excused. This is especially true in common situations that involve an immediate danger to the life or bodily integrity of the person under attack.246

Fletcher rejects the claim that persons who defend themselves against innocent aggressors may typically be excused on the ground that this view implies two conclusions he considers mistaken. The first is that third parties are not justified in intervening in order to help the victims of innocent aggressors by harming the aggressors. The second is that

245. See generally Segev, supra note 5 (elaborating on the relation between justification, rationality and mistake).

246. See Alexander, supra note 241, at 1184-88 (explaining the view that defending oneself against an innocent aggressor or threat is not justified but can often be excused); compare Draper, supra note 66, at 83 (noting that excuse is especially relevant when people defend themselves against immediate threats); Gorr, supra note 6, at 248, 250, 252-53 (claiming that the employment of defensive force against innocent threats is “permissible” but not “completely right”).
innocent aggressors who fight back are justified.\textsuperscript{247} It seems, however, that the first implication is not disturbing, once properly understood, and the second is simply false.

The first implication is not disturbing when we remind ourselves that third parties are justified in intervening in order to stop the attack of innocent aggressors, and thus resolve the conflict, in light of the principles governing the interaction of non-responsible parties. Most notably, intervention is justified if it could resolve the conflict without harming either party. The thesis suggested in this article also justifies intervention by third parties in order to resolve such conflicts when this involves harm to one of the innocent parties (which will typically be the aggressor) provided that the intervention minimizes the harm to both parties to a roughly equal extent or if the harm inflicted to one is significantly less severe than the harm the other party is otherwise expected to suffer.\textsuperscript{248} What is unjustified is to prefer the interests of one innocent party over another on the ground that one constitutes the threat when he is not responsible for this fact according to a choice-based conception of responsibility.

The second implication that Fletcher attributes to the view that people who defend themselves against innocent aggressors are excused - namely, that innocent aggressors who fight back are justified - is simply mistaken. Although a person who defends herself against an innocent aggressor by harming the latter might not be justified, this obviously does not imply that the aggressor is necessarily justified. If a person who defends herself against an innocent aggressor is not justified, since he deviates from the proper principles that govern the interaction of innocent parties, and therefore is, at most, only excused, why is an innocent aggressor, who fights back in a similar manner against a similarly innocent attack, necessarily justified? If the innocent aggressor also deviates from the above principles, then both parties are, at most, only excused.\textsuperscript{249}

\begin{footnotes}
\item[247] Fletcher, Proportionality and the Psychotic Aggressor, supra note 216, at 375.
\item[248] See discussion supra Part III.C.
\item[249] See McMahan, supra note 8, at 286-87.
\end{footnotes}
VI. CONCLUSION

This article's thesis is that the justification for the employment of defensive force, like the justification for the employment of force in general, should be based on the interaction of considerations of well-being and fairness. These considerations include the relative importance of the interests of the persons involved in the conflict and the relative responsibility of the persons for the existence of the conflict. The justification to harm an unjustified and culpable aggressor, in the paradigm case of self-defense, is the most obvious example for the latter consideration.

However, it is important to bear in mind that the paradigm case of self-defense is only one example for the applicability of the general principles suggested in this article for the resolution of interpersonal conflicts, including the fairness-based reason to prefer a person who is not responsible for the existence of a conflict over a person who is, or who is more, responsible. At the same time, it is also important to remember that the implications of the proposed principles might be very different in other kinds of interpersonal conflicts, in which the nature and degree of responsibility for the conflict, and its interaction with other potentially conflicting considerations, might radically differ.250

250. Several other kinds of interpersonal conflicts, which differ, particularly, in the nature and accordingly in the significance of responsibility, are discussed in Segev, Well-Being and Fairness, supra note 9; Segev, Well-Being and Fairness in the Distribution of Scarce Health Resources, supra note 9; and Segev, The Concept of Lesser Evil, supra note 9.