



Are Violations of Rights Ever Right?

Arthur Isak Applebaum

Ethics, Vol. 108, No. 2 (Jan., 1998), 340-366.

Stable URL:

<http://links.jstor.org/sici?sici=0014-1704%28199801%29108%3A2%3C340%3AAVORER%3E2.0.CO%3B2-K>

Ethics is currently published by The University of Chicago Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ucpress.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

Are Violations of Rights Ever Right?*

Arthur Isak Applbaum

If moral constraints (and, correlatively, moral rights) are fundamental and serious elements of morality, can appeals to good consequences justify their violation?¹ Alternatively, if appeals to good consequences sometimes justify the violation of constraints, are constraints on the pursuit of good consequences not fundamental and serious, and so is consequentialism the correct moral theory? This article explores a class of actions where the claim that the actor is morally permitted to violate moral constraints is thought by many to be most plausible, and therefore most challenging to those who want to take constraints seriously. These are cases of *violation minimization*, where the violation of a constraint against one will prevent more numerous or more severe violations of the same constraint against others: violence to the few that prevents bloodier violence to the many, the lie that leads to fewer lies, the coercion that frees others from coercion. The consequentialist argues that if not violating a constraint is good, preventing many violations of the same constraint must be better. The nonconsequentialist replies that that misses the point of constraints, which is to govern how you should treat other persons, despite the consequences. The consequentialist replies: Why should how you treat another have more moral importance than the way many treat many others? And on it goes.²

The permissibility of violation-minimizing violations has been vigorously debated recently because many consequentialists and nonconsequentialists seem to agree that if violating constraints for the sake of

* My interest in this topic was first prompted by Elizabeth Kiss. Deborah Hellman, Stephen Latham, Petr Lom, Martha Minow, Dennis Thompson, Alec Walen, Alan Wertheimer, and Kenneth Winston were generous with written comments. I have learned, though perhaps not enough, from conversations with David Estlund, Frances Kamm, Christine Korsgaard, Arthur Ripstein, and Tamar Schapiro. Anonymous readers and editors of *Ethics* provided exemplary assistance. Work on this article was supported by the Harvard University Program in Ethics and the Professions and the Princeton University Center for Human Values.

1. Roughly, constraints and rights are fundamental if they cannot be reduced to some other moral concept; they are serious to the extent that they have force in moral justification independent of the force of some other moral reason.

2. The problem of violation minimization is first posed, and imaginatively treated, by Robert Nozick, *Anarchy, State and Utopia* (New York: Basic Books, 1974), chap. 3.

Ethics 108 (January 1998): 340–366

© 1998 by The University of Chicago. All rights reserved. 0014-1704/98/0802-0001\$02.00

minimizing violations is morally permissible, then such constraints cannot be fundamental and serious elements of morality. They disagree, of course, about whether such action indeed is permissible. I hope to show that permitting violations under certain circumstances does not undermine the moral point of constraints, and that violating constraints to avoid greater violations is not a contradiction. Why this may be so for violation-minimizing violations of constraints becomes more clear once we look at reasons for permitting the violation of constraints in two related classes of cases: what I shall call *self-defeating constraints*, where not violating a constraint against harming a person will lead to a more severe violation of that constraint against harming that same person, and *Pareto-inferior constraints*, where not violating a constraint against harming a person will not prevent the constraint against harming that person from being violated, and, in addition, will lead to more violations of constraints against harming others.

I. NONVIOLATION AND INVIOABILITY

Frances Kamm and Thomas Nagel have argued, in somewhat different ways, that rights are independently valuable because they express the inviolable status of persons.³ The fact that persons are, are viewed as, and view themselves as the sort of creatures that others must not violate is itself valuable, quite apart from whether persons are actually violated, and quite apart from whether having the status of inviolability reduces the likelihood of being violated. Nagel presents the value of inviolability as a substantive reason why violation-minimizing violations of constraints are ordinarily impermissible. In addition to the substantive reason, Kamm offers a logical one: to violate a constraint against violation for the sake of inviolability involves a contradiction.

Although Kamm's is the earlier account, let us begin with Nagel's.⁴ "Morality is possible," says Nagel, "only for beings capable of seeing themselves as one individual among others more or less similar in general respects—capable, in other words, of seeing themselves as others see them."⁵ Our lives matter to us, but we realize that the lives of billions of others matter to them. Either we concede that, from an impersonal

3. I am grateful to Elizabeth Kiss for introducing me to the arguments about inviolability. See her paper, "Rights and Inviolability: A Critique of Nagel and Kamm's Anti-instrumentalism," which was presented at the eastern division meetings of the American Philosophical Association, 1993; Thomas Nagel, "La valeur de l'invocabilité," *Revue de métaphysique et de morale* 99 (1994): 149–66, portions of which appear in Thomas Nagel, "Personal Rights and Public Space," *Philosophy and Public Affairs* 24 (1995): 83–107, pp. 86–93; F. M. Kamm, "Harming Some to Save Others," *Philosophical Studies* 57 (1989): 227–60, and "Non-consequentialism, the Person as an End-in-Itself, and the Significance of Status," *Philosophy and Public Affairs* 21 (1992): 354–89.

4. Nagel acknowledges his reliance on Kamm in "La valeur de l'invocabilité," pp. 160–63.

5. *Ibid.*, p. 160.

point of view, no one really matters, or we recognize that the fact that each life matters to the one who is living it does matter, in some way, from an impersonal perspective. If one ought to matter to no one but oneself, then one cannot continue to think of oneself as a being that matters very much at all. Since we think of ourselves as beings that matter, consistency demands that we extend that status to others. Says Nagel, "I believe, as did Kant, that what drives us in the direction of universalizability is the difficulty each person has in regarding himself as having value only *for himself*, but not *in himself*. If people are not ends in themselves—i.e. impersonally valuable—then they have a much lower order of worth."⁶ If one wishes to view and value oneself as a being that is an end in itself, and not as a means to be used for the ends of others, then the status of an end must be extended to others. The violation of other persons—using them as means—therefore is an impersonal bad, something we all have reason to avoid and prevent.

But if such violation is bad, why should we not seek to minimize violations, even if that sometimes requires a lesser violation? Because a violation-minimizing violation uses one as a means for the ends of others and so fails to treat persons as ends in themselves. If persons are to matter in the highest possible way, then morality must value not only the absence of violations of persons, but the treatment of persons as beings who have the status of being inviolable—whose violation is not permissible. "What actually happens to us is not the only thing we care about: What *may* be done to us is also important, quite apart from whether or not it *is* done to us—and the same is true of what we *may do* as opposed to what we actually do."⁷ Since having the status of inviolability is of great value, if morality permits violations so as to maximize the good of not being violated, all persons cease to have a high degree of inviolability, which is a great bad. We all may be better off in a world in which morality always treated us as ends, and so where it is always morally impermissible to violate us, even though we are thereby more likely to suffer violation at the hands of immoral actors.

We might think at this point that Nagel is simply adding another variable to the consequentialist calculation, weighing the value of having the status of inviolability against the value of actual nonviolation. If that were so, constraints and rights would not be serious and fundamental after all, but rather, rules of thumb summarizing the trade-off between two more fundamental values of similar consequential structure. Nagel invites this interpretation, for he speaks of one value outweighing the other, and he believes that there are thresholds beyond which the right of inviolability must give way to the consequence of re-

6. *Ibid.*, p. 161.

7. Nagel, "La valeur de l'invocabilité," p. 157, "Personal Rights," p. 91.

ducing violations.⁸ To prevent a large enough number of killings, one is permitted to kill. "Even if rights are not instrumental, they are not absolute. Their extent depends in part on other values besides the non-instrumental value of inviolability."⁹

Let us set aside what Nagel means by 'noninstrumental' and what he thinks has that attribute, the right not to be violated or the value of inviolability. Nagel holds that the status of high inviolability is the outcome of universalization, rather than a maximization: by asking what ways of being treated by others are reasonable to want, we arrive at how others should be treated by us.¹⁰ Nonviolation and inviolability are not, in the first instance, possibly competing values to be weighed across persons from some impersonal perspective—that *would* entail a consequentialist maximization. Rather, nonviolation and inviolability are weighed in the first instance within the life of a person who values them because his life matters to *him*—and this result is universalized to others. This opens the possibility that a person can imagine circumstances under which his own violation ought to be permitted. Therein lies part of the solution to the puzzle of violation minimization . . . but we are getting ahead of ourselves.

Nagel borrows this distinction between inviolability as a status to be respected and nonviolation as a consequence to bring about from Kamm. Both acknowledge the substantive value of treating persons as creatures that have a high degree of inviolability by virtue of being ends in themselves.¹¹ But Kamm goes further, arguing that violating constraints for the sake of minimizing violations of constraints involves confusion and contradiction. Kamm asks:

Why is it not permissible, indeed obligatory, for us to do this—i.e., engage in what has been referred to as a utilitarianism of rights—as an expression of respect and concern for rights?

8. Nagel, "La valeur de l'inviolabilité," p. 164.

9. *Ibid.*, p. 165.

10. *Ibid.*, p. 162. The substantive content of the outcome of some such process of universalization may be "maximize good consequences," but the justificatory force of such content comes from the fact that it is the result of universalization. See, for example, T. M. Scanlon, "Contractualism and Utilitarianism," in *Utilitarianism and Beyond*, ed. Amartya Sen and Bernard Williams (Cambridge: Cambridge University Press, 1982), pp. 103–28, 110, 115 ff.

11. In an important difference, which will not be explored here, Nagel believes that noninstrumental rights cannot be explained without invoking agent-relative reasons and values, while Kamm holds that an entirely agent-neutral, "victim-focused" account of rights is both sufficient and correct. See Thomas Nagel, *The View from Nowhere* (New York: Oxford University Press, 1986); Kamm, "Harming Some to Save Others," pp. 244, 249–51, 255, and "Non-consequentialism," pp. 382, 385–86. For another argument against Nagel, see Christine M. Korsgaard, "The Reasons We Can Share: An Attack on the Distinction between Agent-Relative and Agent-Neutral Values," *Social Philosophy and Policy* 10 (1993): 24–51.

I suggest, in answer to this question, that a moral system—where a moral system is our attempt to represent moral truth—that permits minimization of the violations of a certain right by transgression of that very right essentially eliminates that right from the system, hence it would be futile as a way of showing respect for rights: it would be a “futilitarianism” of rights. This is so, at least, if the right is what I shall call a *specified* right, and rights with at least some sort of specification, we shall see, are the serious rights.¹²

Kamm believes that it would be confused to violate a specified right—a constraint that is not vague about the extent of protection it accords—to minimize violations of that right. For example, suppose (as Nagel allows) there is some threshold beyond which the harms of violation do outweigh the value of inviolability, so that the correct specification of the right not to be killed is: “Do not kill one to prevent the killing of fewer than six others.” Further suppose that others will violate this constraint five times, by killing five persons without justification, unless you violate the constraint once, by killing one to save the five. To violate the constraint for its own sake under such circumstances is confused, says Kamm, because it denies that “Do not kill one to prevent the killing of fewer than six others” is the correct specification of the constraint against killing. “Suppose people are covered by such constraints whose ground is the expression of personal inviolability of a certain sort. Then it would be simply self-contradictory for it to be morally permissible to minimize violations of the constraint itself for the sake of showing concern for *it*.”¹³

II. VIOLATING PERSONS AND VIOLATING RIGHTS

At bottom, I believe, Kamm and Nagel’s arguments traffic in an ambiguity about the subject of nonviolation or inviolability. Are persons not to be violated and treated as inviolable, or are rights not to be violated and treated as inviolable? Kamm and Nagel distinguish, appropriately, nonviolation from inviolability, and see them as separate values. If we also distinguish persons from rights, we have four concepts and four possible values: the nonviolation of persons and the inviolability of persons, the nonviolation of rights and the inviolability of rights.

What do I mean by person violation? Semantically, rape is the clearest case: to rape a woman is to violate her, quite apart from violating her right not to be violated. The concept of person violation extends readily to other ways of damaging bodily integrity: stabbing, torturing, starving, chaining. But not all violations of bodily integrity are person violations,

12. Kamm, “Harming Some to Save Others,” p. 252. “Utilitarianism of rights” is Nozick’s coinage; see Nozick, p. 28.

13. Kamm, “Non-consequentialism,” p. 384.

and not all person violations are violations of bodily integrity. Surgery involves penetration of the body, is painful, and can be disfiguring, but consensual surgery is not a person violation. Suppression of the freedom of thought and discussion is a person violation, though it involves no bodily harm. Let us say, provisionally, that a person is violated when her moral agency, and so her personhood, is either damaged or denied. One's moral agency is damaged when one's capacity to choose and pursue ends is impaired, as when one is killed, crippled, or severely deprived and degraded. One's moral agency is denied when one is treated merely as a tool for the ends of others, as when one is forced, coerced, or deceived. What distinguishes the scalpel from the stiletto, of course, is that the scalpel ordinarily is wielded for an agency-preserving end that the patient, exercising agency, has willed as her own. For now, we will have to make do with this suggestive but sketchy account of person violation. What is to count as damage to or denial of personhood, and so, what is to count as person violation, will become more clear as we progress.

To violate a person is usually to do something terrible, which is why violations of persons are usually violations of the rights of those persons not to be violated. But the two notions of violating a person and violating a person's right not to be violated are separate, and some steps are needed to get from one to the other. Not all violations of persons are violations of rights: striking a willful attacker in self-defense is a person violation, but not a rights violation. To use Mill's famed example, forcibly restraining a man about to walk unknowingly onto a hazardous bridge also violates the person, but no right of the person.¹⁴ Nor are all violations of rights violations of persons. I can wrongfully use your property without using you, and I can damage your property without damaging your capacity for moral agency.

Person inviolability, Kamm and Nagel argue, is a status persons have by virtue of being ends in themselves. This status is expressed and its value recognized through constraints against person violation, and correlative rights not to have one's person violated. The nonviolation of rights is of great moral importance because it realizes the value of not violating persons and expresses the value of the inviolability of persons. But the concept of nonviolation of rights and its value do not collapse into either the concept of nonviolation of persons and its value or the concept of the inviolability of persons and its value.

Because rights nonviolation is of great moral importance, it is a state of affairs to bring about, both through not violating rights oneself and by preventing others from violating rights. Posed this way, of course, we do not yet have an answer to the puzzle of rights maximization, for sometimes the way to prevent others from violating rights is to violate

14. John Stuart Mill, *On Liberty*, ed. David Spitz (1859; reprint, New York: Norton, 1975), chap. 5, p. 89.

rights. For Kamm and Nagel to be correct about the impermissibility of violating rights to minimize their violation, *rights* must have the status of inviolability.

Rights inviolability is a claim about the status of rights, not, at least not directly, about the status of persons. Now, rights must claim some measure of inviolability if rights are to be reasons for action with any sort of prescriptive force. But why this is so, and why rights are to be taken seriously, cannot be because rights are ends in themselves. That is a fetishistic claim, both untrue and unnecessary for taking rights seriously. Persons are ends in themselves, and so have a high degree of inviolability, expressed by rights to nonviolation, and these rights have moral force—that is, some measure of inviolability. But it does not immediately follow from the recognition of a high degree of person inviolability that rights should have an equally high degree of inviolability—that is, that rights should be absolute, or nearly absolute, even when specified. Here, then, is a preliminary answer to Kamm: though indeed it may be a contradiction to violate a constraint in order to express the high inviolable status of constraints, constraints do not have the status of high inviolability. It is no contradiction (though it may be a mistake) to violate a constraint in order to express the inviolable status of persons. The primary objects of our respect and concern are persons, not rights, so there is neither futility nor contradiction in violating the rights of persons if that is what circumstances demand to treat them as ends in themselves. Let us see if there can be any circumstances where this is so.

The problem of rights maximization is one of a family of cases where more or less violation, either of persons or of rights, will occur, depending on what an actor chooses. Through action or inaction, actor A can bring about one of two courses of events. In one, call it the *lesser violation prong*, actor A violates target T. In the other, call it the *greater violation prong*, actor or actors A' violate more severely, more frequently, or in greater numbers target or targets T'. We have distinguished person violation from rights violation, so the violations that are to occur in the lesser and greater violation prongs may be either violations of persons or of rights, yielding the four violation types illustrated in figure 1. Next, we

		LESSER VIOLATION PRONG	
		A violates the person or right of T	
		PERSON	RIGHT
GREATER VIOLATION PRONG	PERSON	I	II
A' violates the person or right of T'	RIGHT	III	IV

FIG. 1.—Violation types

ACTOR(S) A'	TARGET(S) T'		
	T' is T	T' includes T	T' excludes T
A' is A	1	4	7
A' includes A	2	5	8
A' excludes A	3	6	9

FIG. 2.—Identity types in the greater violation prong

specify the identities of actors and targets in both prongs. Target T' in the greater violation prong may be identical to target T, may include target T among other targets, or may be different from target T. Similarly, actor A' in the greater violation prong may be identical to actor A, may include actor A among other actors, or may be different from actor A in the lesser violation prong, yielding the nine identity types illustrated in figure 2. For example, in type 9, actor A faces the decision to violate target T in order to prevent some other actor or actors, A', from causing greater violation to some other target or targets T'. If the violations in both prongs are rights violations (fig. 1, Type IV), then the case described is the standard problem of *rights-violation minimization*, of violating a constraint against one in order to prevent worse constraint violations by others against others. This is the case in which the choice of the lesser violation prong is contradictory, according to Kamm.

Note, however, that this is but one of twenty-eight possible combinations of the four violation types (fig. 1) and the nine identity types (fig. 2). The reader will be spared an elaboration of all twenty-eight, since each is not uniquely interesting, but a few are, and shed light on our question.

III. SELF-DEFEATING CONSTRAINTS

Consider the cases of *self-defeating nonviolation*, where not violating a target will lead to a more severe violation of that same target—that is, where T and T' are identical. We begin with an obvious example, where A and A' are also identical, and the violations in both prongs are person violations, but not rights violations (Type I-1). Suppose that I am attacked by someone who wishes to kill me, and I can defend myself using either the lethal force of a gun or the nonlethal force of my fists. Either way, I do not violate the rights of the attacker, but either way, I violate the person of the attacker. Using lethal force is a greater violation of the person of the attacker. May I take the lesser violation prong and use nonlethal force? Of course I may, and perhaps I must, if I can defend myself just as

well. Though striking a person is a violation, killing a person is a greater violation. I am the one who is harming another in either case. To use a gun in order to avoid using fists is bizarre; surely I may choose the lesser person violation.

Now, a less obvious example: the same attacker menaces me, I have no gun, but I see that an onlooker does, and if I do not defend myself with my fists, the onlooker will shoot the attacker in my defense (Type I-3).¹⁵ Again, I certainly am permitted to punch the attacker. He has no right not to be punched. Though I violate his person, I spare him a graver violation: being shot. If I punch him, I do so for the sake of his personhood: to minimize the violation of *him*. Acting for the sake of another's personhood is not simply to act in another's interests. Rather, it is to respect and preserve, insofar as the circumstances allow, his moral agency. Here, of course, I do not share the attacker's end to do me harm, and in using violence to thwart his will, I violate him; but in sparing him a graver violation, I both respect an end I reasonably assume he does will, not to die, and in so doing I preserve his capacity for moral agency. If person violation is permitted in simple cases of self-defense, a fortiori it is permitted here, where to refrain from person violation is self-defeating.¹⁶

Another example of the same type: the horn of the rhinoceros fetches an astronomical price on the black market because in parts of Asia it is believed to be a powerful aphrodisiac. So conservationists in African game parks saw off rhino horns to prevent poachers from killing the endangered creatures. Now, a rhino is not the sort of creature that has rights, but is the sort of creature that can be harmed, so that the mutilation involved in having its horn sawn off is a "creature violation," a bad to avoid and prevent. But if the conservationists do not inflict this lesser violation, the poachers will commit a far greater violation of the

15. I suppose here that the onlooker is permitted to defend me. If she is not, then this is a case of Type III-3.

16. I note in passing the type variations in which A' includes A, or where A' excludes but is more numerous than A. Cases of these types are useful for exploring our intuitions about various agent-centered reasons and about constraints as objects of concern in themselves. Suppose many actors are to violate the right of a target. Is it morally better that only one actor commit the same violation of the same target, and is that a reason for the one to do so? If so, is this because the number of constraint violators morally matters, or because the number of constraint violations morally matters? It is true that moral wrongness is a nonconservative, nonzero sum property, so that if three lie to another, each of the three has told one full lie, not a third of a lie each, and the constraint against lying has been broken three times. Still, these are cases that point out the overriding moral importance of the wrong done to the patient, rather than of the wrong done by the agent, or the wrong done to constraints themselves, so to speak. It would be very strange to favor an action that involved more severe violation of the target in order to reduce the number of liars or the number of broken constraints against lying.

same rhinos. If the reason one ordinarily should not saw off a rhino's horn is that rhino violation is bad for rhinos, not to saw off the horns in this case for that reason is self-defeating, for the outcome of not sawing off the horns will be much worse for the same rhinos.

A harder case: in Imperial Russia, young boys were sometimes impressed into the tsar's army for many years of harsh, cruel, and dangerous service. Some parents, to spare their sons this fate, would cut off the boys' trigger fingers at a very young age. Unlike rhinos, boys do have rights, and to cut off a boy's finger under the circumstances described is to violate a right not to have one's person violated. Let us suppose that the conditions of conscription amounted to involuntary servitude to a tyrant, and so was a greater violation of these boys' rights. This is a case of Type IV-3, for the parents violate a lesser right of their child to prevent a greater violation by the tsar of the rights of the same child. Such a desperate choice is ghastly to contemplate, and I cannot say with confidence that it is morally permissible. But if the act of mutilation is morally wrong, it is not because the parents fail to respect their child's personhood, insofar as circumstances allow. The child is too young to give genuine consent to either fate. The parents violate his rights for the sake of his personhood, to prevent the tsar both from denying his status as a moral agent by treating him like a tool and from damaging his capacity for agency through deprivation and degradation.

A still harder case: you are a military commander fighting a just war facing the problem of civilian casualties. Success in your campaign requires that you direct bombers to destroy enemy artillery batteries on a nearby hilltop that have been shelling your troops and causing heavy losses. The enemy guns are placed in between a number of civilian houses in a small village on the hill. Hoping to deter your attack, the enemy army has falsely assured the villagers that they are safe. They cling to their homes, some paralyzed by fear, some manipulated into wishful thinking by their army. Your pilots are equipped with smart bombs that can pick off with precision the artillery pieces wedged in between the houses without in themselves causing much other damage, but the bombing is certain to ignite enemy ammunition stores that will blow up and destroy the houses and their inhabitants.

On most accounts of justice in war, you do not violate a moral constraint against violating the civilian noncombatants by bombing the enemy artillery. Either the killing of the civilians is an unintended but foreseen side effect, or the killings are causally more distant than the direct effect, and so the civilians are not being treated as means; the goal of destroying the guns is itself justified, and the civilian losses are proportionate to the gains. Believing that you have no better options, you are about to order the bombing, when an aide who studied moral philosophy before the war proposes an alternative: first aim machine gun fire

directly at the houses in the hope that the civilians will be terrorized into fleeing. There is some risk that you will wound or kill the residents, but the alternative is certain death.

This is a case of Type II-1: the actor's choice is between causing greater person violation without violating a moral constraint or causing the same target lesser person violation by violating a moral constraint. If you shoot directly at the civilians, you do so for the sake of their personhood, not as a means to some other's end. Under the circumstances, to violate the constraint against shooting at them is in the service of their status as inviolable creatures. Given that they are otherwise permissibly doomed, to violate their rights is to treat them, as much as possible, as ends in themselves. Again, the question of consent arises. Do we not treat them in the most respectful way by respecting their choice to perish with their homes? By assumption, they have not genuinely chosen to do so. Their will is already undermined by their army's false assurances. In all the cases we have examined so far, some defect in the moral agency of the target makes nonviolation as a form of respect self-defeating. The attacker has an evil will; the rhino, no rational will; the child, an immature will; the villagers, impaired wills.

Kamm appears to endorse a permission in cases like Type II-1. After proposing a new formulation of the moral constraint on harming some to save others,¹⁷ she briefly introduces an exception called the "Principle of Secondary Permissibility."¹⁸ If an actor is permitted (under Kamm's formulation of permissible harms) to cause harm to a victim, and is physically able to do so, then the actor is also permitted to substitute a lesser harm that would otherwise be prohibited. Kamm does not elaborate on the reason that permits such substitutions other than to say that it is in the best interests of the person who will be harmed. But why should one's interest in minimizing the violations of one's own person be a good enough reason to limit the scope of the constraints that express that person's inviolability? The reason, I believe, is that an unlimited constraint in this instance is self-defeating: its observance would not be for the person's sake, but for the constraint's sake, and so would fail to treat the person as an end in himself.

If this explanation accounts for our judgment in the bombing case, then the explanation extends beyond the exception specified in Kamm's Principle of Secondary Permissibility. What if actor A', who is to inflict the greater but permissible person violation, is not the same as actor A, who can prevent the greater person violation by violating a constraint? Suppose the bombing case were a case of Type II-3; actor A commands the machine gun battery, but has no control over the decision to bomb,

17. Roughly, the harm produced may not be causally closer to the act that causes both harm and benefit than the benefit produced.

18. Kamm, "Harming Some to Save Others," p. 248.

and so cannot physically cause the greater permissible harm. On the irrevocable orders of air force commander A', bombing will begin at a certain hour. If A violates the constraint against the intentional and direct targeting of civilian noncombatants by shooting at the houses before the bombing begins, he will prevent the greater but permissible person violation that A' will cause. Why should it morally matter that the preventable but permissible violation is not to be caused by A? An expansive constraint against person violation is self-defeating in both cases in the same way: observing the constraint fails to treat the victims as ends. Indeed, when A and A' are different, A treats T more as an end than in Kamm's case, where T is in harm's way because A is aiming in the first instance at the good of others. Here, sparing T the greater violation is the only reason A has for action.

Following the reasoning of self-defeating rights, once the greater person violation is to be caused by another (A and A' are different), why should it matter that the greater violation prong is not also a violation of rights? In the tsar's army case (Type IV-3), the parents have a stronger moral reason to mutilate their child, not a weaker one, because military service violates their child's rights, and is not merely a grave but permissible person violation.¹⁹ When T and T' are the same, minimizing rights violations recognizes the inviolability of persons in the highest possible way. Minimizing violations within a life avoids the objections against minimizing violations across lives.

Let us take stock before continuing: by exploring cases of self-defeating nonviolation, we have driven a wedge between constraints and their point and shown that, when the point of constraints is undermined, one morally may, and perhaps must, attend to the point, not the constraint. This is not, however, a victory for consequentialism, for the point of constraints is to treat the person as an end, not the good as an end.

One can, alternatively, hold that no constraint is broken when the lesser violation prong is permitted. Rather, we recognize a new, narrower specification of the constraint. Moral rights can either be understood to be broad in scope but overrideable in the face of countervailing moral reasons—that is, not inviolable—or they may be understood to be narrowly and perhaps elaborately specified in anticipation of all possible countervailing reasons, so that when the right so specified applies, it applies absolutely, without exception, and so is inviolable. Which is the

19. It is not generally the case, however, that the lesser harm is easier to justify when the greater harm is brought about by a rights violation, rather than by a permissible person violation or a natural misfortune, because one has a reason not to become a tool of evil. It may matter whether the lesser violation resists the greater violator, as in the tsar's army case, or collaborates with the greater violator, as in the Jim and the Indians case, below. For a discussion of these matters, see N. Ann Davis, "The Priority of Avoiding Harm," in *Killing and Letting Die*, 2d ed., ed. Bonnie Steinbock and Alastair Norcross (New York: Fordham University Press, 1994), pp. 335–44.

right way to think about rights, and moral principles generally, is a question too large to tackle here, but one comment is in order. For those who believe that the possibility of intrinsic, noninstrumental rights depends on their inviolability, it is an empty victory if the only way for constraints to hold absolutely is to specify their scope with excessive narrowness. Again, the point of moral constraints must be kept in full view: treating persons as ends by recognizing that they have the status of creatures with a high degree of inviolability. Moves to ensure that *rights* have a high degree of inviolability, by building into them inviolability as a formal condition, may not serve to treat persons as creatures with a high degree of inviolability. If the two ways of thinking about rights yield identical answers about when, all things considered, one is justified in violating a person, there is no reason to believe that the scheme that treats rights as inviolable treats persons as more inviolable, and there is a reason to believe that the opposite is true. One whose right has been justifiably violated may have a residual moral claim or complaint against the violator, and the violator may incur subsequent moral obligations to the violated.²⁰ One who rightly violates another's right has reason for reluctance and regret, and perhaps even remorse and resentment. Morality's recognition of these remainders reaffirms the moral importance of the victim even as it permits victimization.

IV. PARETO-INFERIOR CONSTRAINTS

Consider now those cases, call them *Pareto-inferior constraints*, where not violating a constraint against violating a person will not prevent the constraint from being violated, and, in addition, will lead to more violations of constraints against violating others. The much-studied, nightmarish story of Jim and the Indians offered by Bernard Williams is a case. Jim comes across twenty natives picked at random who are about to be executed to discourage political protest. The captain in charge offers Jim, as an honored foreign guest, the privilege of shooting one native, in which case the nineteen would be set free. The Indian who would be chosen by Jim has a right not to be killed, but will be killed in any case—will not merely die, but will be killed in violation of his right in any event. This is a case of Type IV-6: Jim, actor A, must choose between the lesser rights violation prong, in which he kills target T, or the greater rights violation prong, in which the captain's henchman, actor A', kills the twenty targets T', including T among them.

Jim and the Indians is a case of Pareto-inferior rights because if Jim observes the right of one Indian, that Indian is no better off, and nine-

20. Compare the common law of necessity: I am permitted to use any port in a storm, but if I cause damage to your dock, I am liable. I find it implausible to say that in such situations one is merely excused, not fully justified, as if to drown and not trespass is a supererogatory action, rather than a neurotic one. But one still is responsible for the consequences.

teen Indians are worse off. This is so, not only in the space of well-being, but also in the space of rights: the right of the one is violated just the same, and nineteen additional rights violations occur. Unlike in the cases of self-defeating rights, where A violates the rights of T for the sake of the personhood of T, here A (Jim) violates the rights of T (one Indian) for the sake of nineteen others. But does that mean that Jim fails to treat the one that he kills as an end in himself, thereby denying that the one has a high degree of inviolability?

Suppose it is like this: the Indian to be shot by Jim is to be chosen at random, and Jim's decision to participate must be made before the lottery. Or suppose Jim is handed a rifle with nineteen blanks and one live round. If Jim can find the appropriate *ex ante* place to stand, he can understand what he does as subjecting each to a one-in-twenty risk of death in order to spare each certain death. More important, the Indians can understand Jim this way. Posed like so, this becomes a case of a self-defeating right: each, *ex ante*, faces the lesser violation at Jim's hand. He threatens each with the risk of death for the sake of each. *Ex ante*, all are violated less. *Ex post*, none is violated more. If Jim shoots, he does not obviously deny the point of rights.

Now, the next step: all cases of Pareto-inferior rights are cases of self-defeating rights from some suitably constructed *ex ante* perspective. In the face of unknown threats to one's moral agency over the course of a lifetime, a person generally is accorded more protection from those threats, and in no particular case is accorded less protection, by a morality that permits the violation of Pareto-inferior rights. If a general principle sometimes is to a person's advantage and never is to that person's disadvantage, then actors who are guided by that principle can be understood to act for the sake of that person. It would be self-defeating if, to respect and not deny that person's moral agency, an actor were guided by an alternative principle that never damaged that person's capacity for moral agency less, and sometimes damaged it more.

But why should moral reasoning proceed from the *ex ante* perspective? There are many arguments for adopting many variations of the *ex ante* perspective, of course, and some version of the *ex ante* point of view has been offered to represent, model, or constitute a number of related but distinct moral ideas: impartiality, reciprocity, universality, generality, the rule of law, autonomy, and various forms of consent. Though this discussion will concentrate on the connection between the *ex ante* perspective and varieties of consent, I will not argue here for a favored formulation of the *ex ante* perspective, or for the primacy of any one of the moral ideas the *ex ante* perspective is said to explicate. Rather, I simply wish to point out one appealing feature of the generic *ex ante* strategy, which is that our imaginative ability to hold in our heads one version or another of the *ex ante* point of view while we deliberate about how to act *ex post* helps to explain why cases like Jim and the Indians perplex us.

When we reason from both perspectives we see how a violation of a self-defeating or a Pareto-inferior constraint can both affirm the inviolable status of persons and still remain a genuine violation. This is so because there are some ways of being treated to which consent is impossible, *ex post*, but to which consent is both possible and reasonable, *ex ante*.

V. POSSIBLE CONSENT AND REASONABLE CONSENT

Let us begin with the Kantian claim that some ways of being treated are impossible to consent to, so that when one is treated in those ways, one necessarily is used as a means. Christine Korsgaard explains:

Kant's criterion most obviously rules out actions which depend on force, coercion, or deception for their nature, for it is of the essence of such actions that they make it impossible for their victims to consent. If I am forced, I have no chance to consent. If I am deceived, I don't know what I am consenting to. If I am coerced, my consent itself is forced by means I would reject. So if an action depends upon force or deception or coercion, it is impossible for me to consent to it. To treat someone as an end, by contrast, is to respect his right to use his own reason to determine whether and how he will contribute to what happens.²¹

If persons are violated when they are treated as mere means, and if persons are treated as mere means when they are treated in ways to which they cannot possibly consent, then, straightforwardly, persons are violated when they are treated in ways to which they cannot possibly consent. The argument runs in two directions. The forward claim: some ways of treatment, such as deception and coercion, by their nature undermine or overpower the will, and so cannot be the object of genuine consent. The backward claim: if, by hypothesis, one *has* given genuine consent to some way of being treated, then such treatment cannot have been deceptive or coercive.

The criterion of possible consent has been read by some very broadly, so that whenever I cannot veto an action that significantly affects me, either because I do not have the knowledge or do not have the power to refuse treatment, I am treated as a means.²² Posed so, the crite-

21. Korsgaard, "The Reasons We Can Share," p. 46 (footnote omitted). See also Christine M. Korsgaard, "The Right to Lie: Kant on Dealing with Evil," *Philosophy and Public Affairs* 15 (1986): 325-49, p. 332; Onora O'Neill, "Between Consenting Adults," in her *Constructions of Reason* (Cambridge: Cambridge University Press, 1989), p. 113.

22. The broad view is held by O'Neill: "But those closely involved in or affected by a proposal have no genuine possibility of dissent unless they can avert or modify the action by withholding consent and collaboration. If those closely affected have the possibility of dissent, they will be able to require an initiator of action either to modify the action or to desist or to override the dissent. But an initiator who presses on in the face of actively expressed dissent undercuts any genuine possibility of refusing the proposal and chooses rather to enforce it on others." See O'Neill, "Between Consenting Adults," pp. 110-11; see

tion is extremely far reaching. For instance, you apparently treat me as a means on this view if you seek to compete with my business in a way that would be economically harmful to me without giving me effective power to veto your plan. Without taking a stand on whether the broad reading is the best reading of Kant, I will construe the criterion more narrowly here in order to distinguish harmful actions that use the will of others as a tool (deception, coercion, and torture) from harmful actions that simply do not defer to the will of others (damaging property). Actions that fall under the narrow criterion most clearly involve treatment as means, and most straightforwardly fail the possible consent test, since the central aim of such actions is to subvert the will of the victim. With regard to actions of the second sort, there of course is a sense in which actions that affect another fall under two different act descriptions, depending on whether the person affected has the genuine opportunity and capacity to veto the treatment or does not. It indeed is impossible, in that sense, for you to consent to the act described as "damaging your property without giving you the opportunity to refuse," just as it is obviously impossible for you to consent to the act described as "damaging your property without your consent." Were you to give genuine consent to the property damage, you would be consenting to a differently described action. But there is no redundancy in the description "damaging your property without your consent" as there is in the description "coercing you without your consent," and there is no internal contradiction in the description "damaging your property with your consent," as there appears to be in "coercing you with your consent." I will focus on the clear, core instances of impossible consent.

In instances where consent is impossible, it makes no sense to ask if consent is reasonable. Even if reasonable consent is the correct test of moral permissibility, it can be applied only to act descriptions to which consent is a possibility. Genuine consent to a particular act of coercion is a conceptual impossibility, for genuine consent renders the act something other than coercion. What, then, could be meant by asking if it is reasonable to consent to a particular act of coercion? How can one have reason to do what is conceptually impossible? Possible consent is a precondition for reasonable consent. So, even if the correct test of moral permissibility is reasonable consent, insofar as violations are actions to which consent is impossible, the notion of a permissible violation does not get off the ground.

One way to get the notion of a permissible violation going is to show that persons and their rights can be violated even when they can consent to their treatment. The forward claim, that some ways of treatment can-

also p. 139, n. 12, and p. 217. Korsgaard supposes that consent is possible as long as the target has "some power over the proceedings" (Korsgaard, "The Right to Lie," p. 333), but not necessarily a veto.

not be the object of genuine consent, does not hold for the sorts of violation that are not as intimately tied to the undermining or overpowering of the will as are deception and coercion. It is conceptually possible to consent to be maimed or killed. Therefore, for some sorts of violation, the backward claim also fails: having consented to have one's person violated by maiming or killing, such treatment ordinarily remains the person violation, and also the rights violation, of maiming or killing. There are exceptions, where consent changes not only the prescription, but the description of the action: surgery certainly is not battery, voluntary euthanasia probably is not murder, and we need not describe either as a person violation. But the violation involved in most instances of wounding and killing survives consent. I am inclined to think that prizefights are morally impermissible because fighters aim at and inflict grave person violation on each other. But there is nothing about the concepts of consent and violence that renders consensual violence, and therefore, a consensual prizefight, a conceptual contradiction.

Willing treatment that aims at the destruction of one's capacity for moral agency, such as killing, crippling, or extreme degradation, may involve a contradiction in the will, and so involve a kind of irrationality.²³ This is a reason not to accord such consent the moral force consent usually has to render actions permissible. But it does not follow that irrational but nonetheless genuine consent by an otherwise competent, free, and informed person is impossible. So, consent to at least some violations is possible, without eliminating the violation. Sometimes, such consent justifies the violation, though the violation remains a violation; and sometimes, such consent does not justify the violation, though the violation remains consensual. Whether actual consent permits violation in these cases could turn on whether such consent is reasonable, for, where consent is possible, there is no conceptual block to the application of a test of reasonable consent. And if that is so, then there is no block to applying a test of reasonable consent to possibly consensual, but actually nonconsensual, violations. Though we should expect most of such violations to fail such a test, some could pass.

VI. EX ANTE CONSENT THAT IS IMPOSSIBLE TO ENDORSE EX POST

Can the notion of a permissible violation get off the ground in the harder and more interesting case, where consent to the violation appears to be impossible? If reasonable consent is the test of moral permissibility,

23. See Onora O'Neill, "Consistency in Action," in her *Constructions of Reason*, pp. 89 ff., for the notion of volitional inconsistency without universalizing. See also Barbara Herman, "Murder and Mayhem," in her *The Practice of Moral Judgment* (Cambridge, Mass.: Harvard University Press, 1993), pp. 118 ff., who argues that—Kant's treatment of the suicide example notwithstanding—a universalized maxim of killing fails only the contradiction of will test, not the contradiction in conception test.

we seem to be in a bind, for the question of reasonableness does not arise where consent is an impossibility. The way out is to note that the reasonable consent test can be applied, not to particular acts, but to general practices and policies. Even for those treatments that are intimately tied to undermining the will, such as deception, there is no conceptual impossibility in actually consenting in advance to a policy of deception, even though it remains impossible to consent to each particular incidence of deception. For example, one can consent to a central bank's general practice of lying about interest rate hikes, or the police's general practice of undercover sting operations, and still be successfully deceived in the particular case.²⁴ It can be both possible to will the general practice, *ex ante*, and impossible to will the particular application, *ex post*. Whether or not such deception is justified by general consent, general consent is at least possible, though the treatment remains deceptive.²⁵

The claim that it is possible to will the general practice *ex ante* but impossible to will the particular application *ex post* is open to two obvious objections, and answering them will help to clarify the claim. The first objection is that the possibility of *ex ante* consent entails the possibility of *ex post* consent. Yes, it is possible to give consent to a general practice of deception, but in so doing, one renders the particular instances of deception that fall under the practice consensual via some distributive property that consent possesses. To consent *ex ante* to the general practice simply is to consent to the particular acts *ex post*. Deception, it turns out, is not really different from violence: consensual deception is no contradiction. With deception, we need to interpose some uncertainty-creating distance to allow for deception's success, and this does not need to be done in the case of violence, but once this gap is opened, consensual deception is possible.

Alternatively, one may object that, for other ways of being treated, neither *ex ante* nor *ex post* consent is possible. Consider force. If consent to a general practice renders particular instances consensual, then one cannot give genuine consent to the practice without thereby giving genuine consent to the particular acts. But if one has given genuine con-

24. See Dennis F. Thompson, *Political Ethics and Public Office* (Cambridge, Mass.: Harvard University Press, 1987), pp. 22–24, 26–29, for a discussion of general consent.

25. Korsgaard seems to hold that, while one successfully can consent in advance to be left uninformed, examples of consent in advance to be deceived that succeed in deceiving are more difficult to describe. See Korsgaard, "The Right to Lie," p. 333, n. 6. While this may be so about advance consent to a one-shot lie, in the case of advance consent to a general practice of deception, successful deception is not at all difficult to describe, because of the increased complexity of inferring that you are in a situation where lying is to be suspected. Indeed, Korsgaard herself describes a case where a publicized general practice of deception *can* succeed: she supposes that Kant's murderer at the door mistakenly thinks that his intentions are hidden, and so does not suspect that he will be deceived, even though he knows about the general practice of lying to murderers. In any case, merely withholding information can itself be a person violation.

sent to the particular acts, then what one has given consent to cannot be described as force, because consensual force is a contradiction. And if the particular instances cannot be described as force, then it would not be a general practice of force that one was consenting to, but a general practice of something else.

This objection can be posed another way. For consent to be genuine, it must be free and informed. For general consent to distribute genuine consent over particular instances, and so be genuine general consent, one must be free enough and informed enough with respect to particular treatments up front. But the standard of free enough and informed enough is quite stringent. Because filling out information about future deceptions in a way sufficient to satisfy the informed consent condition undermines the possibility of success of future deceptions, one cannot give genuine consent to a general practice of deception. Similarly, one cannot give genuine general consent to be forced, for there is a defect in the voluntariness of consent when one agrees to unspecified future forcings. For the consent to be truly free, one would need to give consent freely to each of the future forcings. This requires that, in the future, one will have a morally acceptable alternative to each proposed treatment.²⁶ But if this is so, each particular treatment is genuinely voluntary, not forced. Informed consent to a general practice of deception cannot succeed, and free consent to a general practice of force lacks sense.

These objections call for refinements. Insofar as the first objection is on target, clarify the claim like so: it is possible to give general *ex ante* consent to some ways of being treated for which consent would be impossible to *initiate* or *endorse* *ex post*. In situations where consent cannot possibly be initiated, we recognize a danger of person violation, because one cannot at the time of treatment endorse one's prior willingness. This is so, not simply because one does not have the contingent opportunity to endorse, as the wide view of impossibility would have it. If I give you general permission to reveal my confidences to others and it turns out that I am not present when you do so in particular, the person violation of betrayal does not threaten. But this is because there is no conceptual difficulty with endorsing permission to tell my secrets at the time of the telling. Success in the act of revealing my confidences does not require that I be unable to endorse the act. In granting you permission in advance to reveal confidences, I run the risk that your future treatment of me will be contrary to my future will. But ordinarily, I can exercise that will in some way, either by withdrawing the granted permission in situations where I have the authority to do so, or, where the permission is not

26. Or perhaps it is sufficient that at the time general consent is given one expects to have a morally acceptable alternative in each *ex post* instance. But this is quite different from a much weaker test, that at the time general consent is given one has a morally acceptable alternative to the expectation of all the *ex post* instances—that is to say, an acceptable alternative to the *risk* of harmful treatment.

unilaterally revocable, by petitioning you not to exercise it. You need not accept such a petition, but it would be indecent to reject it without offering reasons that appeal to my moral agency. And in offering reasons, even if only to say "Stop whining, we made a deal, I did my part, now do yours," you treat me as a person, not as a tool. In contrast, when you rely on prior general consent to deceive me, you aim to undermine my future willing in a way that prevents me from genuinely endorsing my prior willing, and prevents you from offering me reasons in the face of my reluctance. Similarly, if you rely on prior general consent to force me, you plan to overpower any future reluctance in a way that utterly disregards the question of endorsement. The practical success of deception and the conceptual coherence of force depend upon the local denial of my moral agency.

Ex post unwillingness by itself does not render harmful treatment wrong, of course. You may have the moral right or even the obligation to act on the strength of the prior consent, and this may be so even when the victim looks you in the eye and pleads, "Stop!" But our sensibilities in such a situation are clear: we are inflicting harm on an unwilling victim, and, in some cases, this sensibility gives rise to a sense: a person is violated, even if no right of the person is violated. I think this is why capital punishment is indecent whether or not it is unjust.²⁷ Some such sense of violation may be behind the common law principle that specific performance of a contract is rarely compelled, though an unwilling contractor is liable for the losses caused by nonperformance.

As the distance increases between the point of view of one who grants prior general consent to be harmed and the point of view of one subjected to harmful treatment, subsequent endorsement cannot be taken for granted. Minds change over time, of course. But by distance between points of view, I do not mean the passage of time only. The greatest distance between the ex ante and ex post perspective is that between the risk of some harmful treatment and its certainty. The difference between the two kinds of distance is the difference between a Faustian bargain and a Shylockian one.

The response to the second objection, that both ex ante and ex post consent are impossible, makes use of this second kind of distance, the distance between consenting to risks of treatment and consenting to certain treatment. For the familiar reasons, one cannot initiate genuine consent to a certain and immediate instance of force without rendering the treatment unforced. The novice skydiver, frozen in place, who asks his instructor to push him out of the plane is not forced. But one can genuinely agree, in advance, to the risk of force without rendering the ex post

27. See Avishai Margalit, *The Decent Society* (Cambridge, Mass.: Harvard University Press, 1996), pp. 280–81 for the notion of a just but indecent society, and pp. 290–91 for the notion of sense and sensibility in moral concepts.

treatment unforced. A borrower who agrees to indentured servitude in the event of default is, upon default, forced.

The position stated in the objection, that one cannot give free and informed consent to the risk of harm if one does not have both full knowledge of and an acceptable alternative to each certain harm, is false. If it were true, then much of the vast network of agreements we engage in to pursue our plans and projects over time in coordination with others would be defective. To give informed and free consent to a risk, one need only be informed and free with respect to the risk. To be informed about a risk one needs rough probabilities and magnitudes, but not prophecy, and to freely accept a risk one needs to have a morally acceptable alternative up front to entering into an agreement under uncertainty up front. One does not need an acceptable alternative *ex post* to each harmful eventuality.²⁸

Now, there is a perfectly understandable sense in which consent to the risk of harm distributes over the realized instances. If I place a bet on the horses and lose my stake, at first look it seems bizarre for me to claim that I consented only to the risk of losing, but did not consent to actually losing. Someone who sincerely made such a claim would appear not to understand what a bet is. But it does not seem bizarre at all to say that one who consented to the risk of dying in the boxing ring did not consent to actually die in the boxing ring. It surely is not the case that the boxer wills to die. The boxer, knowing the risks (and rewards), wills to box.²⁹ Why can we not say that the bettor, knowing the risks and rewards, wills to bet? It would be bizarre to claim that *ex ante* consent to risk does not ordinarily distribute the relevant moral permissions, obligations, liabilities, or immunities over the *ex post* realizations. But it is not bizarre to maintain that one who wills an agreement does not will each contingency under it, Rousseau's well-known argument notwithstanding. Rousseau says about giving up one's life in defense of the state: "The purpose of the social treaty is the preservation of the contracting parties. He who wills an end wills the means to that end: and the means in this case necessarily involves some risk, and even some loss. He who wills that his life may be preserved at the expense of others must also, when necessary, give his life for their sake."³⁰ But he who wills a means does not necessarily will all of its consequences. If Rousseau counts the

28. I mean to leave open just how to evaluate the moral acceptability of an agreement under uncertainty and its alternatives. Expected utility theory is but one of several ways.

29. Compare Kant on punishment: "No one suffers punishment because he has willed *it* but because he has willed a *punishable action*; for it is no punishment if what is done to someone is what he wills, and it is impossible to *will* to be punished." Immanuel Kant, *The Metaphysics of Morals*, trans. Mary Gregor (1797; reprint, Cambridge: Cambridge University Press, 1991), pp. 143–44 (Prussian Academy edition, p. 335).

30. Jean-Jacques Rousseau, *The Social Contract*, trans. Christopher Betts (Oxford: Oxford University Press, 1994), bk. 2, chap. 5, p. 71.

fatal outcome as a means to the end of self-preservation, he is mistaken. The assumption of the risk of death is the necessary means; actual death is a stochastically foreseeable but unwilled consequence. Whether one nonetheless is obligated to give one's life is another matter.

It follows that one can willingly accept the risk of a harmful treatment that one could not willingly accept *ex post* without changing the description of the treatment. Therefore, one coherently can consent to the risk of force or coercion and still describe the realized treatment as force or coercion. As long as the treatment cannot be genuinely endorsed when it happens, it remains without contradiction a forcing, and so a presumptive violation. Therefore, it is possible to give genuine *ex ante* consent to risks of violations that are impossible to endorse, *ex post*.

About what is possible, we can ask what is reasonable. If one can possibly consent to a violation at the level of *ex ante* general practice, we are not precluded from asking what moral principles governing violation are reasonable to consent to. The answer may indeed be that it is never reasonable to consent to certain sorts of violation, but need not be. A moral principle that permits violating persons when not to do so is Pareto-inferior, but that nonetheless preserves the sense that a person has been violated, cannot be ruled out on conceptual grounds alone, for it involves no contradiction.

Grounding principles by appeal to some version of reasonable agreement is a major enterprise in contemporary moral and political philosophy, of course, and it is not my purpose here to add yet another friendly amendment to the projects of Rawls or of Scanlon. I simply wish to point out that contractualist accounts in general can make the following two claims: first, they can admit the moral permissibility of intentional violation in some cases without undermining or contradicting their commitment to the inviolability of persons, and so without conceding that rights are merely instrumental. Second, they can do so without having to redescribe the violation as something else.

VII. THE POSSIBILITY OF PERMISSIBLE VIOLATION

For concreteness, consider one such view. I believe that the argument I have made here is roughly compatible with Thomas Hill's Kantian account of when doing "the lesser evil" is justified, though he reaches somewhat different conclusions.³¹ Hill argues that the appropriate point

31. See Thomas E. Hill, Jr., "Moral Purity and the Lesser Evil," in his *Autonomy and Self-Respect* (Cambridge: Cambridge University Press, 1991), pp. 67–84, and "Making Exceptions without Abandoning the Principle: Or How a Kantian Might Think about Terrorism," in his *Dignity and Practical Reason* (Ithaca, N.Y.: Cornell University Press, 1992), pp. 196–225. "Moral Purity" is largely concerned with showing why principles that *require* doing the lesser evil in some cases would fail to be adopted by ideal legislators. "Making Exceptions" aims to show why principles that *permit* doing the lesser evil in some cases might succeed in being legislated.

of view from which to formulate moral rules is that of rational, impartial legislators who share a full range of moral ends and values, but are not yet committed to specific rules of conduct, and who are aware that not all will comply with the rules that they enact.³² From that point of view, says Hill, "one might argue, all would favor any policy that promised to prolong the survival of representative persons as rational autonomous persons, other things being equal and provided that the policy is otherwise consistent with the dignity principle. Each legislator would favor the policy when looking at it from his or her own perspective (abstractly conceived) and also when reviewing it from the perspective of each other person."³³

Hill concludes that a policy of rescuing otherwise-doomed hostages that will foreseeably but unintentionally endanger some in attempting to save all is not inconsistent with the dignity principle, and therefore could be agreed upon from the general legislative point of view. Hill therefore permits actions that cause violation as an unintended side effect where not to cause violation is, in the terms we have been using, self-defeating to the violated. Hill does not explicitly address the unintentional Pareto-inferior case, where, though the prospects of those who are foreseeably endangered are neither improved nor worsened, others will benefit from the rescue attempt. But, a fortiori, we can infer his stand from his views on a tougher case. Even when the innocents who would foreseeably be endangered are not already endangered, Hill holds that a carefully circumscribed permission "could be defended from the general legislative point of view even to those who turn out to be the victims under the policy."³⁴ If the policy increases the chances of all to live full lives as rational agents, and if all have roughly the same chance of becoming victims, a permission to endanger some bystanders in order to rescue many hostages could conceivably be consistent with human dignity. And if unintentionally causing the violation of one who would not otherwise be violated in order to minimize the violation of many others might be justified, surely unintentionally causing the violation of one already fated to be violated—the Pareto-inferior case—might be justified.

Hill draws a clear line between merely foreseeable and intentional harming, however, and holds that the deliberate killing of someone already doomed could not be justified from the legislative point of view: "Thus, to be justified in a deliberate killing, a person would need to be

32. Hill, "Moral Purity," p. 79. In an older essay, Hill puts it like this: "To adopt our principles as ideal legislators seems a good idea; but to make them for ideal law-followers does not" ("The Kingdom of Ends," in his *Dignity and Practical Reason*, p. 66). See Korsgaard, "The Right to Lie," for a different way of finding room for nonideal theory in Kantianism.

33. Hill, "Making Exceptions," p. 215. The dignity principle refers to Kant's second formula, that humanity in each person be treated as an end, not only as a means.

34. *Ibid.*, p. 216.

able to face the victim and say, sincerely and truthfully, 'I choose to kill you (when I have an option not to); but still I regard you as more than a mere means, in fact, as a person with a worth that is incalculable.' Now, it is hard to imagine that a human being could maintain this attitude while deliberately killing another, even though apparently Kant thought that public executioners could and many think that doctors administering euthanasia can.³⁵ But this judgment about the moral importance of deliberateness does not follow inexorably from Hill's own argument. Though the thrust of his work is to show that Kantianism is not as rigoristic as is often thought, I think that here he retains more rigorism than his account requires. For impartial legislators reasoning about what principles to follow in the face of evil do not obviously devalue or disrespect each other if they agree to save each other from violation at the hands of evildoers by subjecting each other to lesser or less likely deliberate violation at the hands of fellow moral citizens. To repeat an earlier proposition: if a general principle sometimes is to a person's advantage and never is to that person's disadvantage, then actors who are guided by that principle can be understood to act for the sake of that person. It would be self-defeating if, to respect and not deny that person's moral agency, an actor were guided by an alternative principle that never damaged that person's capacity for moral agency less, and sometimes damaged it more. It is not unreasonable for impartial moral legislators to permit the deliberate violation of constraints when the constraints are self-defeating, and such permission need not be interpreted as maximizing nonviolation across persons or putting a price on human dignity, for the permission is for the sake of each person, *ex ante*. *Ex post*, one could say to the particular victim for whom the constraint is no longer self-defeating, but simply Pareto-inferior, "A general policy of permitting deliberate violation in this situation would have been adopted for your sake, precisely because you are a person of incalculable worth. It would not have been unreasonable for you to have agreed to it, *ex ante*, and for you to withhold agreement now, *ex post*, is futile." Something like this, I have claimed, is how Jim can understand what he does if he shoots one Indian to save nineteen, and how both the one and the nineteen can understand that what is done to them by Jim is done *for* them.

Something like this, but not quite this—for though we can imagine the Indians understanding what Jim does, genuinely agreeing or withholding agreement to their treatment is not a possibility for them. Against the background threat of murder at the hands of the captain, there is no plan that Jim can propose to which the Indians can give uncoerced consent. It is not that Jim coerces them, of course—the captain does. But the Indians cannot freely consent to Jim's actions. This is in part why, if Jim shoots, he commits a grave violation, however the In-

35. *Ibid.*, p. 220 (qualifying footnotes omitted).

dians understand what he does. In a suitably constructed *ex ante* choice situation, it is not unreasonable for persons who put the highest value on each other's humanity to agree to the violation of self-defeating and Pareto-inferior constraints—that is what might make such violations permissible. But we must not lose the sense that what is permitted is still, *ex post*, a genuine grave violation. Part of what makes the killing of the one Indian a genuine violation is that, against a coercive background, no one can possibly give genuine consent to such treatment.

This is not to say that what the Indians want Jim to do is morally irrelevant. On Williams's telling, the Indians not only can understand what Jim does, but do understand: they are begging Jim to accept the captain's offer. Korsgaard, in an illuminating discussion of this example, points out the importance of what the Indians in fact want Jim to do. It matters, she says, whether the Indians are pacifists who plead with Jim not to participate in the captain's violence, or whether the oldest Indian steps forward and volunteers his life. "*Very* roughly speaking, you are not treating him as a mere means if he consents to what you are doing. Of course, the Indian does not in general consent to be shot, and his gesture does not mean that after all he has not been wronged. In the larger moral world he has. But if you and the Indians are forced to regard Pedro and the captain as mere forces of nature, as you are in this case, then there is a smaller moral world within which the issue is between you and them, and in that world this Indian consents."³⁶ Korsgaard qualifies this by warning of the spuriousness of simply saying yes, and cites with approval O'Neill's test of having an authentic opportunity to say no. So Korsgaard apparently believes that, within this "smaller moral world," it is conceivable that the Indians can be given an authentic opportunity to refuse.

Now, the main point seems right—the way for Jim to treat the humanity of the Indians with as much respect as the horribly unjust situation allows is to let them decide their own fate, insofar as they are able. But one might be tempted to draw three mistaken inferences from this: that consent by itself is generally sufficient to permit violation, that consent always is necessary, and that the consent of Indians in the smaller moral world renders Jim's shooting something other than a violation. Consent is not always sufficient: if Jim ought to defer to the wishes of the Indians, it is because a policy of letting the victims decide insofar as they are able in such situations satisfies a test of reasonable agreement, and the futility of their refusing to be violated makes a decision favoring violation reasonable. If some weird but sane thrill seeker under no threat of death genuinely requested Jim to shoot him, Jim must refuse, because permission to aid in such self-degradation would not pass Hill's or any

36. Korsgaard, "The Reasons We Can Share," p. 46 (qualifying footnote omitted). In Williams's story, Pedro is the captain's gunman.

other test of reasonable agreement among person-valuing persons. Consent is not always necessary: if Jim cannot ask the Indians how they wish to be treated, he is not obviously required to let them all die. And, even though the Indians consent in the smaller moral world to risk being shot by Jim, Jim still violates the one he shoots. This is obscured by the example of the elder who volunteers to be sacrificed, for from his extraordinarily courageous standpoint the captain's threat does not crush all meaningful choice. He has a salient and reasonable alternative, a lottery that gives him a nineteen in twenty chance of surviving, and he chooses not to propose it. But suppose instead that none of the twenty Indians is as brave and selfless as the elder. The lottery is proposed, and all agree—after, let us say, each has been given by Jim the opportunity to opt out and be shot for sure by the captain. The choice looks a lot less free now, even in the “smaller moral world,” for the only alternative each has is certain death. Straws are drawn. If the loser loses his nerve and pleads with Jim to be spared, I am not sure that the best response is, “In the smaller moral world you genuinely consented.” The doomed man quite rightly replies, “But what choice did I have?” A better response appeals to what is reasonable for all *ex ante* and fair to the others *ex post*, rather than to an agreement forced by necessity. If the loser sees a way to escape both Jim and the captain, I am not sure that he is obligated to stick around. If he is, fairness, not consent, is the source of the obligation. If Jim is permitted to shoot him in the back as he runs for safety, the justification is not *volenti non fit injuria*: the Indian has not, in any meaningful sense, willed away the constraint against his own violation. And if this is so for the unlucky coward, it is also so for the unlucky stoic who accepts the results of the lottery and faces Jim's bullet with equanimity. *Ex post* endorsement of a coerced *ex ante* agreement is still unfree. The point still stands: when constraints against person violation are self-defeating or Pareto-inferior, what might render breaking these constraints permissible is that such treatment passes a test of reasonable *ex ante* consent. What might keep such treatment a violation is that it fails the test of possible *ex post* endorsement.

Kamm ends her discussion of harming some to save others with the suggestion that one might permissibly bargain away the status of inviolability for the outcome of fewer violations, but she sees such agreements as belonging to a different moral topic, that of permissible bargains.³⁷ If, instead, we try to represent what reasonable contractors might find reasonable to agree to in the face of unreasonable circumstances, we may find that the topic is the same.

Is one ever justified in violating the rights of some to prevent the greater violation of the rights of others? We can now see that one need not em-

37. Kamm, “Harming Some to Save Others,” p. 256.

brace a "utilitarianism of rights" to answer yes. A moral theory that views persons as intrinsically valuable does not require that constraints be treated as ends in themselves, or that they have a high degree of inviolability. Rather, persons ideally are to be treated as ends in themselves, and understood to have a high degree of inviolability, which is expressed by constraints on how they may be treated. If constraints are understood to have a high degree of inviolability, nothing follows about the treatment of persons, for constraints that have the formal property of inviolability might be extremely limited in scope.

How can sacrificing the rights of one for another be compatible with highly valuing each person as an end in itself with a life to lead? The answer has both a formal and substantive part. Formally, anything that a person reasonably would accept in a suitably constructed position of choice is compatible with highly valuing persons. As a substantive matter, what sorts of sacrifices are reasonable to accept? In the face of injustice, sacrifices for the sake of establishing the principle that persons are of high intrinsic value strike me as plausible candidates for reasonable acceptance—more plausible than sacrifices for the sake of maximizing good consequences simply. Rawls distinguishes ideal theory—the principles of justice that hold in a well-ordered society, when all are in compliance with them—from nonideal theory—the principles that hold in the face of injustice and unfavorable social and economic conditions that stand in the way of a well-ordered society.³⁸ The principles of nonideal theory adjust to the realities of injustice and misfortune, so that the ideal of treating others as ends sometimes gives way to the goal of a world in which others are treated as ends. Because the goal is justified by a test of reasonable acceptance in the face of unreasonable circumstances, not by its good consequences, nonideal theory does not, I believe, collapse into consequentialism. We may disagree about whether any cases of violating the rights of some to prevent the greater violation of the rights of others meet such a test of reasonable acceptance. But there is no conceptual difficulty in entertaining that some do.

38. John Rawls, *A Theory of Justice*, secs. 2, 11, 26, 39, 46, 53–59. See Korsgaard, "The Right to Lie," for a sketch of a Kantian nonideal theory.