



## Self-Defense

Judith Jarvis Thomson

*Philosophy and Public Affairs*, Vol. 20, No. 4 (Autumn, 1991), 283-310.

Stable URL:

<http://links.jstor.org/sici?sici=0048-3915%28199123%2920%3A4%3C283%3AS%3E2.0.CO%3B2-B>

---

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.

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.

*Philosophy and Public Affairs* is published by Princeton University Press. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/pup.html>.

---

*Philosophy and Public Affairs*  
©1991 Princeton University Press

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact [jstor-info@umich.edu](mailto:jstor-info@umich.edu).

©2003 JSTOR

I

Before we take a close look at it, we may think of self-defense as morally transparent. What could be clearer than that morality permits a person to save his or her life against threats to it?

But what if in order to save one's life one has to kill another person? In some cases that is obviously permissible. In a case I will call *Villainous Aggressor*, you are standing in a meadow, innocently minding your own business, and a truck suddenly heads toward you. You try to sidestep the truck, but it turns as you turn. Now you can see the driver: he is a man you know has long hated you. What to do? You cannot outrun the truck. Fortunately, this is not pure nightmare: you just happen to have an antitank gun with you, and can blow up the truck. Of course, if you do this you will kill the driver, but that does not matter: it is morally permissible for you to blow up the truck, driver and all, in defense of your life.

It is probably not necessary to stress *here*—though the point is important for our purposes—that you do not merely have an excuse for blowing up the truck, you are morally permitted to do so. Suppose Peter killed Paul. For it to be the case that Peter merely has an excuse for killing Paul is for it to be the case that though he ought not have done so, and acted wrongly in doing so, still he is not as much at fault for doing so as he would have been had he not had the excuse, and, if his excuse is not merely mitigating but completely exculpating, then he is not at fault at all for doing so. But blowing up the truck in *Villainous Aggressor* is not something you ought not do. We cannot plausibly say that you ought not

An early version of this article was presented at the University of Arizona Philosophy Colloquium; I am grateful to those present for their comments. A number of people made helpful comments on a later draft: I am particularly grateful to Jonathan Bennett, Bruce Ackerman, and the Editors of *Philosophy & Public Affairs*.

blow up the truck, but will only be in a measure at fault, or in no measure at fault, for doing so: you simply *may* blow up the truck. Morality permits it.

But why may you blow up the truck? The following answer suggests itself: the driver is villainously aggressing against you, and will thereby kill you unless you stop him—and, moreover, you can stop him only by blowing up the truck. That second clause is important. Suppose you had two ways by which you could stop him. One, blow up the truck with your antitank gun, or two, wave your antitank gun and shout, “Stop, or I’ll blow up the truck!” If you do not actually need to blow up the truck, if you really could stop him by merely threatening to blow up the truck, then that is what you ought to do. It would be wrong to kill even a villainous aggressor when you do not need to do so.

So the driver is villainously aggressing against you, and will thereby kill you unless you stop him, and you can stop him only by blowing up the truck. Why not say that *that* premise explains why you may proceed? I suggest that it is in at least one way stronger than it need have been.

## II

For let us look at a second hypothetical case, which is like the first except in this respect: the driver is entirely without fault for what he is doing. How can that be, given that he is chasing you around the meadow in a truck, trying to run you down? Well, let’s suppose some villain had just injected him with a drug that made him go temporarily crazy. It is not his fault that he is going to kill you if you do not blow up the truck, he is not villainously aggressing against you; but he is aggressing against you, and he will in fact kill you if you do not blow up the truck. Does morality permit you to blow up the truck? *I* think it does: I think self-defense permissible in this case—which I will call Innocent Aggressor—just as in the case I called Villainous Aggressor.<sup>1</sup>

1. Innocent Aggressor is modeled on George P. Fletcher’s splendid hypothetical case in “Proportionality and the Psychotic Aggressor,” *Israel Law Review* 8 (1973): 367–90: “Imagine that your companion in an elevator goes berserk and attacks you with a knife. . . .”

I will everywhere be using “innocent” to mean free of fault. But it might pay to take note here of another use of “innocent”—sometimes “technically innocent”—under which it means “not harming”: see G.E.M. Anscombe, “Mr. Truman’s Degree,” reprinted in her

But I suspect that some people would say that while self-defense is permissible in Villainous Aggressor it is merely excusable in Innocent Aggressor. Should we agree with them? By hypothesis, the driver in Villainous Aggressor is villainously aggressing against you and the driver in Innocent Aggressor is not; but how exactly could that difference be thought to show that while it is permissible for you to proceed in Villainous Aggressor, it is merely excusable for you to proceed in Innocent Aggressor?

Perhaps it will be thought that the point is this: the villainous driver is less worthy (less good) than the fault-free driver. But the fault-free driver too might be a not particularly worthy person. His being fault-free in Innocent Aggressor consists only in his being free of fault for the particular aggression that he is currently committing.

Alternatively, perhaps it will be thought that the point is this: the villainous driver deserves punishment for his aggression, whereas the fault-free driver does not. But who are you, private person that you are, to be dishing out punishment to the villainous for the things that they do? And anyway, what makes it permissible for you to blow up the truck in Villainous Aggressor is not the fact that the driver in that case deserves punishment for what he is doing, or else it would be permissible for you to blow up the truck even if you do not need to do so to save your life.

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

It might help also if I stress that I am not saying here that the fault of an aggressor is nowhere relevant to the question what you may do to defend yourself. I say here only that it is not relevant when what is in question is his life for yours. What if we lower the cost to you? I should think that fault is also irrelevant when the aggressor would otherwise blind you, or cut off your legs: the aggressor's fault or lack of fault has no bearing on whether you may kill the aggressor to defend your eyes or legs. (Death is worse than blindness or being without legs, but other

things being equal, morality does not require submission to such an aggressor, even if he is without fault.) I should think, more generally, that the same holds whenever the aggressor would otherwise cause you very grave bodily harm. Similarly if we go to the other extreme and imagine the prospective cost to you is minimal. If the aggressor would otherwise take your wallet or hat, then you may not kill the aggressor to defend yourself, whether or not the aggressor is at fault for his aggression. (That it is a villain who will otherwise take your wallet or hat is not sufficient reason to kill him to protect it.) There is room for argument, however, at places between these extremes of very grave bodily harm on the one hand, and loss of wallet or hat on the other hand. Suppose an aggressor will take, not both your legs, but only your left foot unless you kill him. Here the aggressor's fault or lack of fault may well be thought to make a difference: thus it may be thought that you may kill him to defend your left foot against his aggression if he is at fault but not if he is without fault. I leave this open. (I also leave open what should be said in cases in which it is not certain that the aggressor will cause you a harm if you do not kill him but only more or less probable that he will.) What I think is clear in any event is that if the aggressor will (certainly) take your life unless you kill him, then his being or not being at fault for his aggression is irrelevant to the question whether you may kill him.

But as I said, I suspect that some people would say that self-defense is merely excusable in Innocent Aggressor—or at least would feel uncomfortable about saying that it is permissible. I have a hypothesis as to what is at work in anyone of whom this is true, and will come back to it later.

Suppose, however, that we are in agreement that morality permits blowing up the truck in Innocent Aggressor as in Villainous Aggressor, so that you do not in either case merely have an excuse for proceeding. Then we are in agreement that the fault of the driver in Villainous Aggressor does no moral work in making it permissible for you to proceed in that case. But then the premise we looked at, by appeal to which it might be thought possible to explain why you may proceed in Villainous Aggressor, was stronger than it need have been. (I fancy we overrate the role of fault in many areas of moral theory; further examples will turn up shortly.)<sup>2</sup>

2. I think we also misinterpret the role played by fault. See, e.g., my "Morality and Bad Luck," *Metaphilosophy* 20 (1989): 203–21.

Then why may you blow up the truck in the two cases? Why not just say this: the drivers are aggressing against you—in the one case villainously, in the other case without fault—and will thereby kill you unless you stop them, and you can stop them only by blowing up the trucks, and *that* premise explains why you may proceed.

But isn't this premise, too, stronger than it need have been?

### III

In a third hypothetical case—which I will call Innocent Threat—you are lying in the sun on your deck.<sup>3</sup> Up in the cliff-top park above your house, a fat man is sitting on a bench, eating a picnic lunch. A villain now pushes the fat man off the cliff down toward you. If you do nothing, the fat man will fall on you, and be safe. But he is *very* fat, so if he falls on you, he will squash you flat and thereby kill you. What alternative do you have? Well, you only have time to shift the position of your awning; if you do this, the fat man will be deflected away from you. But deflecting him away from you will be deflecting him past the edge of the deck down onto the road below. Does morality permit you to shift the awning? I think it does.

As I said in the preceding section, I suspect that some people would say that while self-defense is permissible in Villainous Aggressor it is merely excusable in Innocent Aggressor; I suspect that even more people would say that self-defense is merely excusable in Innocent Threat. Why so? The innocent aggressor, though without fault, is at least aggressing against you; the fat man is not only without fault, he is not doing anything at all—he is merely *falling* toward you.

I think that difference makes no moral difference, and thus that it is permissible for you to proceed in Innocent Threat just as in Villainous Aggressor and Innocent Aggressor. I have a hypothesis as to what is at work in those who think otherwise, and will come back to it later.

Suppose we are in agreement that morality permits proceeding in Innocent Threat just as in Villainous Aggressor and Innocent Aggressor, so that you do not in any of these cases merely have an excuse for proceeding. Then we are in agreement that the aggression of the drivers in

3. Similar cases have been discussed by others, but unless I am mistaken, the first to draw moral philosophers' attention to them was Robert Nozick, in *Anarchy, State, and Utopia* (New York: Basic Books, 1974), pp. 34–35.

Villainous Aggressor and Innocent Aggressor does no moral work in making it permissible for you to proceed in those cases. But then the premise we looked at, by appeal to which it might be thought possible to explain why you may proceed in those cases, was stronger than it need have been.

How to weaken it? The premise says: the drivers in Villainous Aggressor and Innocent Aggressor are aggressing against you and will thereby kill you unless you stop them, and you can stop them only by blowing up the trucks. It follows that the drivers in those two cases will kill you unless you kill them. Can we say of the fat man in Innocent Threat that he too will kill you if you do not kill him? If so, then it is plausible to think that what is at work in this third case is the same as what is at work in the first two: that is, it is plausible to think we can explain the permissibility of proceeding in all three cases by appeal to the fact that in all three, the man you kill if you proceed *will otherwise kill you*.

It might be replied, however, that this is not true of the fat man in Innocent Threat. I said earlier that if he falls on you, he will squash you flat and thereby kill you, but is it really right to say that if he falls on you he will *kill* you? It might be said that while an event consisting in his fall on you will kill you, *he* will not have killed you. After all, he will not have aggressed against you, he will merely have fallen on you.

If we are moved by that idea, then we will think that the premise should not be weakened in the way I just pointed to. Perhaps, then, we will have the following thought: even if it is not true of all three—the two drivers and the fat man—that they will otherwise kill you, they are anyway all threats to your life, threats to your life that you can defend against only by killing them. And perhaps we will think *that* is what the premise should be weakened to.

I think that we should not be moved by that idea. Is it really to be thought that Y kills X only if Y aggresses against X? Suppose a piano and a safe fell off a roof, and we know that one fell on Alfred, and that the event that consisted in its fall on him killed him. We might ask, “Which killed Alfred, the piano or the safe?” The correct answer might be, “The piano,” despite the fact that pianos commit no acts of aggression. (“Which of these bullets killed Kennedy?” asks the museum curator of the bullets just donated to the museum by the police. “That one,” say the police, pointing to one in particular. And they might be right.) Indeed, I

should think that if an event that consists in the fall of Y on X kills X, then it follows that Y killed X, whatever Y may be.

Moreover, the suggested alternative weakening of the premise is not self-explanatory. What exactly is it for one thing to be a threat to the life of another? The two drivers will kill you if you do not kill them, and perhaps that clearly enough marks them as in some appropriate sense threats to your life. If we say it is not true of the fat man that he will kill you if you do not kill him, then what exactly does he have in common with the two drivers that is supposed to mark him too as in that same sense a threat to your life?

So I suggest we reject both the idea and the suggested alternative weakening of the premise; I suggest we say that, like the drivers in *Villainous Aggressor* and *Innocent Aggressor*, the fat man in *Innocent Threat* will kill you if you do not kill him. There is no need to be confused: we need merely remember about the fat man that, though he will kill you if you do not kill him, he will not have killed you by an act of aggression. If we agree to this, and I will assume we do, then we can say that in all three cases, the person you kill if you proceed *will otherwise kill you*.

And then if we are in agreement that morality permits proceeding in all three cases, why not say it is that premise that explains why?

Well, it will not do. I think we are helped to get at what is missing if we attend first to some cases in which the premise is false.

#### IV

The three cases we have been looking at are cases in which I think it permissible for you to kill a person in defense of your life; I will therefore call them Yes cases. It is perfectly plain, on any view, that there are No cases, cases in which it is impermissible for you to kill a person in defense of your life. Three classes are of particular interest.

In the first place there are *Substitution-of-a-Bystander* cases. In a case I will call *Trolley*, a villain has started a trolley down a track toward you. You cannot stop the trolley, but you can deflect it. Unfortunately, the only path onto which you can deflect it will take it onto a bystander who cannot get off the path in time. (I intend this allusion to the *Trolley Problem*, and will return to it briefly later.) I take it to be plain that you may



not proceed. If you proceed, you will be making a bystander be a substitute victim; hence the name of the class.

A second class of cases might be called *Use-of-a-Bystander* cases. A villain has started a trolley down a track toward you, and the only way you have of defending yourself is to shoot a bystander who stands on a footpath over the track: he is sufficiently heavy to crush the trolley's rooftop mechanism when he falls onto it, which will thereby stop the trolley. In this case too it is plain that you may not proceed. To proceed would be to use a bystander as a piece of equipment; hence the name of the class.

The mark of a *Use-of-a-Bystander* case is that you need the bystander you would have to kill if you are to defend your life. In a *Use-of-a-Bystander* case, if the bystander goes miraculously out of existence just before you act, then there is no way at all in which you can defend yourself. That is not true in *Substitution-of-a-Bystander* cases such as *Trolley*. But there are cases that fall into both classes, as, for example, where you defend yourself against a villain's gunfire by grabbing a bystander and using him as a shield so that he is shot instead of you. That case is both a *Substitution-of-a-Bystander* case and a *Use-of-a-Bystander* case.

A third class of cases I will call (for want of a better name) *Riding-Roughshod-over-a-Bystander* cases. A villain is shooting at you, and your only defense is to run. But your only path to safety lies across a bridge that will hold only one person, and there is already a man on it; if you rush onto the bridge, he will be toppled off it into the valley below. In this case too it is plain that you may not proceed. If you proceed, you do not make a bystander be a substitute victim, so it is not a *Substitution-of-a-Bystander* case. If you proceed, you do not use a bystander as a piece of equipment, so it is not a *Use-of-a-Bystander* case. What you do if you proceed is to "ride roughshod" over a bystander; hence the name of the class.

Proceeding in some of these cases is worse than proceeding in others. Here is another *Use-of-a-Bystander* case, which I will call *Starvation*. Some villains throw you in a dungeon, and leave you there, without food, for several weeks. Not surprisingly, you are by now very hungry. To tempt you, and thereby increase your misery, the villains now introduce a plump baby into the dungeon. They remove the baby for feeding periodically, so that *it* is at no risk of starvation. But you are. May you eat the baby? Of course not. It would presumably be far worse to eat the baby in

Starvation, which is a Use-of-a-Bystander case, than it would be to turn the trolley in Trolley, which is a Substitution-of-a-Bystander case. If, in the press of fear, you turned the trolley in Trolley, then we might think this, though wrongful, nevertheless more or less excusable. If, in the press of pain and fear, you ate the baby in Starvation, that would be both wrongful and in no measure excusable.<sup>4</sup> But proceeding in Use-of-a-Bystander cases is not everywhere worse than proceeding in Substitution-of-a-Bystander cases and Riding-Roughshod-over-a-Bystander cases. Suppose you are a subway track workman. A subway is headed toward you. (1) There is a small alcove in the wall near you, but there is another workman already in it. You can pull him out into the path of the subway and get into the alcove yourself. This is a Substitution-of-a-Bystander case. (2) There is a small alcove in the wall near you, but there is another workman already in it. You can force your way into the alcove, thereby crushing him to death. This is a Riding-Roughshod-over-a-Bystander case. (3) There is no alcove, but there is another workman with you on the track. You can shove him into the path of the subway, which will stop it. This is a Use-of-a-Bystander case. I do not think that there is any difference in the degree of moral badness of your proceeding according as we supply ending (1), (2), or (3) to the story that begins with a subway headed toward you.

Why, after all, should it have been thought that the fact that you *need* the person you would have to kill in order to defend yourself makes it worse for you to proceed than it would have been had you not needed the person? If I am right in thinking that is the mark of a Use-of-a-Bystander case, then *using* a person does not in general have the special moral taint that has been ascribed to it. Appeals to the notion “respect for persons” will certainly not suffice to make out this special moral taint. After all, if one proceeds in a Substitution-of-a-Bystander case (a Substitution-of-a-Bystander case that is not also a Use-of-a-Bystander case) or in a Riding-Roughshod-over-a-Bystander case, one behaves as if the person one kills were not there at all—surely no less a display of lack of respect for persons.<sup>5</sup>

4. Or so I think. Some people think it was in a measure excusable (though of course wrongful) for the sailors to eat the cabin boy in *Regina v. Dudley & Stephens* 14 Q.B. 273 (1884). Perhaps what is at work in those people is the thought that it was likely that all, including the cabin boy, would otherwise die. Even so. . . .

5. I am indebted to Jonathan Bennett here.

Let us go back. Cases of these three kinds are No cases. And in all of them our premise is false: it is not true in any of them that the person you have to kill to save your life *will otherwise kill you*. Indeed, that person is, as I have everywhere put it, a bystander.

## V

Is it everywhere impermissible to kill a bystander in defense of one's life? Consider a case that is often discussed in the literature on the Doctrine of Double Effect. In Strategic Bomber, the villainous Bads have invaded noble country Good, and a Good pilot has been ordered to bomb a munitions factory in Bad. Unfortunately, the factory is so situated that if the pilot bombs the factory, some nearby Bad civilians will be killed—let us imagine them to be children, in a nearby children's hospital. Doesn't morality permit the pilot to proceed all the same? But the children are mere bystanders.

In the literature on the Doctrine of Double Effect, Strategic Bomber is often contrasted with Terror Bomber, in which the Good pilot is ordered to bomb, not a Bad munitions factory, but a Bad children's hospital—the point of having him do this is to terrorize the Bad population and thereby get them to sue for peace with the Goods. Friends of the doctrine say it is clear that while the pilot may proceed in Strategic Bomber, he may not proceed in Terror Bomber, and they claim that the Doctrine of Double Effect explains this moral difference between the cases.

Very roughly, the Doctrine of Double Effect says that we may do what will cause a bad outcome in order to cause a good outcome if and only if (1) the good is in appropriate proportion to the bad *and* (2) we do not intend the bad outcome as our means to the good outcome. Thus, for example, the doctrine yields that the Good pilot in Strategic Bomber may bomb the Bad munitions factory in order to cause the Goods to win the war (good outcome), despite the fact that he will thereby cause the deaths of some children (bad outcome), if while he foresees that he will cause the deaths of the children, their deaths are not intended by him as a means to causing the Goods to win the war. By contrast, the doctrine yields that the Good pilot in Terror Bomber may not bomb Bad children in order to cause the Goods to win the war (good outcome) if *he* would be causing the children's deaths (bad outcome) as a means to causing terror and thereby causing the Goods to win the war. Many people take

it to be an attractive feature of this doctrine that it supplies an explanation of the moral difference between the cases.

Much of the literature on the Doctrine of Double Effect consists, on the one hand, in efforts to make it precise in a way that does not make it yield morally implausible results in other pairs of cases to which it is applied, and on the other hand, in efforts to explain why the distinction to which it draws our attention—that between foreseen and intended effects—is morally relevant.<sup>6</sup>

It is a very odd idea, however, that a person's intentions play a role in fixing what he may or may not do. What I have in mind comes out as follows. Suppose a pilot comes to us with a request for advice: "See, we're at war with a villainous country called Bad, and my superiors have ordered me to drop some bombs at Placetown in Bad. Now there's a munitions factory at Placetown, but there's a children's hospital there too. Is it permissible for me to drop the bombs?" And suppose we make the following reply: "Well, it all depends on what your intentions would be in dropping the bombs. If you would be intending to destroy the munitions factory and thereby win the war, merely foreseeing, though not intending, the deaths of the children, then yes, you may drop the bombs. On the other hand, if you would be intending to destroy the children and thereby terrorize the Bads and thereby win the war, merely foreseeing, though not intending, the destruction of the munitions factory, then no, you may not drop the bombs." What a queer performance this would be! Can anyone really think that the pilot should decide whether he may drop the bombs by looking inward for the intention with which he would be dropping them if he dropped them?<sup>7</sup>

Here is Alfred, whose wife is dying, and whose death he wishes to hasten. He buys a certain stuff, thinking it a poison and intending to give it to his wife to hasten her death. Unbeknownst to him, that stuff is the only existing cure for what ails his wife. Is it permissible for Alfred to give it to her? Surely yes. We cannot plausibly think that the fact that if he gives it to her he will give it to her to kill her means that he may not

6. A recent example is Warren S. Quinn, "Actions, Intentions, and Consequences: The Doctrine of Double Effect," *Philosophy & Public Affairs* 18, no. 4 (Fall 1989): 334–51.

7. Jonathan Bennett's criticism of the Doctrine of Double Effect in *Morality and Consequences*, The Tanner Lectures on Human Values III (Salt Lake City: University of Utah Press, 1981) focuses on its friends' use of the notion 'intention', but he also points on pp. 96–98 to the oddity I point to here, as well as to the ideas that issue in two theses I will shortly propose for adoption.

give it to her. (How could *his* having a bad intention make it impermissible for him to do what *she* needs for life?)

The badness of Alfred's intention may incline some people to want to say that while it is permissible for Alfred to give-his-wife-the-stuff, it is impermissible for Alfred to give-his-wife-the-stuff-to-kill-her. (And compare: while it is permissible for the pilot to drop-the-bombs, it is impermissible for the pilot to drop-the-bombs-to-cause-terror.) It would of course be odd to say to Alfred, "You may give your wife the stuff, but you may not give it to her to kill her"; but that it would be odd to say a thing is compatible with the thing's being true. A puzzle lurks here, however. Some people believe that the conjunction of "If X were to do alpha then X would in fact do beta" and "X may not do beta" entails "X may not do alpha." Now by hypothesis, if Alfred were to give-his-wife-the-stuff, he would in fact give-his-wife-the-stuff-to-kill-her. So on that view, if we say that Alfred may not give-his-wife-the-stuff-to-kill-her, we are committed to saying that Alfred may not give-his-wife-the-stuff. I think that view is mistaken, but I take no stand on it here.<sup>8</sup> I here suggest only that *if* we accept it, then we must also accept that Alfred may give-his-wife-the-stuff-to-kill-her. For whatever else we say, we must accept that Alfred may give-his-wife-the-stuff. After all, she needs it for life.

I suggest, quite generally, that we should accept

*The Irrelevance-of-Intention-to-Permissibility Thesis:* It is irrelevant to the question whether X may do alpha what intention X would do alpha with if he or she did it.

Thus in particular, it is irrelevant to the question whether Alfred may give-his-wife-the-stuff what intention he will do it with if he does it. (Indeed, it is also irrelevant to the question whether Alfred may give-his-wife-the-stuff-to-kill-her what intention he will do *that* with if he does it—thus, for example, whether he will give-his-wife-the-stuff-to-kill-her to inherit her estate or to get her early entry into heaven or what you will.)

There is a more general thesis in the offing here that I think we should also accept—indeed, it is the truth of this more general thesis that explains the truth of the less general thesis. If Alfred gives the stuff to his

8. A recent discussion of a cousin of this issue is Michael J. Zimmerman's "Where Did I Go Wrong?" *Philosophical Studies* 59 (1990): 55–77.

wife he will give it to her to kill her. His giving her the stuff to kill her would be his giving it to her with a bad intention, and he will therefore be gravely at fault if he gives it to her with this intention. But that he will be at fault if he gives it to her with this intention does not mean that he may not give it to her.

The point I make here is simply the other side of a coin we took note of earlier. I drew attention to the fact that a man may be without fault though he does something wrongful; what we take note of here is that a man may be at fault though he does something that is not wrongful. (Supposing that the fact that a man would be at fault in doing a thing fixes that it is impermissible for him to do it is yet another example of our overrating of the role of fault in morality.) In short, I think we should accept a thesis that incorporates both sides of this coin, namely,

*The Irrelevance-of-Fault-to-Permissibility Thesis:* It is irrelevant to the question whether X may do alpha whether X would be at fault in doing it.

And the truth of this more general thesis explains the truth of the less general thesis, for if fault is irrelevant to permissibility, then so also is intention. If it is irrelevant to the question whether Alfred may give the stuff to his wife whether Alfred would be at fault if he did it, then it must also be irrelevant to this question what intention he would give her the stuff with if he gave it to her. Alfred's intention is of moral interest only via its being the case that he will be at fault if he acts on it, and therefore his intention has no bearing on the permissibility of his acting if his being at fault itself has no such bearing.

But what matters for our purposes is only the Irrelevance-of-Intention-to-Permissibility Thesis, for if that thesis is true, then an agent's intending or merely foreseeing an effect of his or her action is irrelevant to the question whether he or she may proceed with it, and the Doctrine of Double Effect collapses.<sup>9</sup> (Not because the doctrine is in need of a yet

9. I should add, however, that this is so only if the doctrine is interpreted in the way I indicated. There are other ways of interpreting it. One possibility is to construe the doctrine as concerned, not with intendings, but with sheer causal order; I ignored this possibility in the text above, since I think it pretty obvious that the doctrine so construed has no future at all. A second possibility is to construe the doctrine as concerned with intendings all right, but not with permissibility, rather with fault. I think it possible to interpret G.E.M. Anscombe's remarks on the distinction between foreseen and intended consequences in her article "Modern Moral Philosophy" in this way: she there speaks of the distinction as bear-

more subtle characterization of the relevant intendings, but because it is a product of a misassessment of the role of intendings in morality.)<sup>10</sup> A fortiori, the doctrine cannot be appealed to to explain the moral difference between Strategic Bomber and Terror Bomber.

## VI

What should we think of Strategic Bomber? I think it a murkier case than has usually been thought. I said that the children whom the pilot would kill are mere bystanders. But then why isn't this a No case, a No case falling into the class of Riding-Roughshod-over-a-Bystander cases?

In Trolley, a villain has started a trolley toward you. You can save yourself only if you deflect the trolley, but there is a bystander on the only path you can deflect it onto, and you therefore may not proceed. Consider a variant on Trolley, which I will call Trolley-Preemption. In this case you cannot deflect the trolley at all, you can only fire your antitank gun at it. But there is a bystander standing next to the trolley track, and if you fire your antitank gun, you will blow up the bystander along with the trolley. Can anyone plausibly think it impermissible to deflect the trolley in Trolley but permissible to blow it up in Trolley-Preemption? Hardly. Trolley is a No case falling into the class of Substitution-of-a-Bystander cases; surely Trolley-Preemption is also a No case, but one that falls instead into the class of Riding-Roughshod-over-a-Bystander cases.

But Trolley-Preemption is Strategic Bomber without the war between

---

ing on "responsibility" and "credit." ("Modern Moral Philosophy" is reprinted in her *Collected Philosophical Papers*, vol. 3.) But so construed, the doctrine does not—by hypothesis it does not—have anything to say about permissibility, which is our topic here.

10. I suspect that this misassessment is made only by those who misassess the role of fault in morality, for I suspect that you would think intendings relevant to permissibility only if you thought fault is. But why do people think fault is? There must be something deep behind this idea, since it is so common, and not merely among friends of the Doctrine of Double Effect. Perhaps what is at work in at least some people who harbor this idea is an inward-looking, stoic conception of the limits of morality. If you think morality is wholly a matter of what makes a person be good or bad—if you suppose it is wholly a set of instructions for being virtuous and not being vicious—then it would be no surprise if you thought that one is in breach of morality only if one does a thing the doing of which marks one as so far a bad person, i.e., a thing one is at fault or to blame for doing. (I think this conception of morality could be described as a kind of moral solipsism.) It is a good question what could make morality true if it goes beyond these limits; but perhaps no better than the question what could make it true even if it does not.

the Goods and the Bads: Strategic Bomber differs from Trolley-Preemption only in that the pilot's proceeding in Strategic Bomber is part of a larger process in which the noble Goods are defending themselves in a war being waged against them by the villainous Bads. If we think it permissible for the pilot to proceed in Strategic Bomber, but impermissible for you to proceed in Trolley-Preemption, then we must grant that that is because there is a war on in Strategic Bomber.

How does the fact that there is a war on in Strategic Bomber do the moral work of making it permissible for the pilot to proceed? At least partly, I am sure, by virtue of making it the case that the (long-range) stakes are higher than we have been told about in being told about those children in the hospital next door to the munitions factory.

Not just anything goes in wartime, of course. The fact that it is villains who, in the course of fighting an unjust war against your side, have so arranged things that you can save your life only by killing a person does not by itself make it permissible to kill the person. For example, it would not be permissible for you to eat the baby in Starvation even in mid-war. Moreover, most people think it impermissible for the pilot to proceed in Terror Bomber. (But why, exactly? Suppose a Good pilot bombed a place in Bad that contained both a munitions factory and a children's hospital, and that the Bads therefore sued for peace—not because of the loss of the munitions factory, but because of the loss of the children: the bombing terrorized the Bads, bringing home to them what war was going to have in store for them. It can hardly be thought that the fact that the causal route to the Goods' winning the war passed through Bad terror, rather than through Bad lack of munitions, shows that it was impermissible for the pilot to drop his bombs. It is of course large-scale terror bombing of enemy cities that people have in mind when they think of Terror Bomber; but why exactly is that wrongful? Well, is it always wrongful? Must it be? Perhaps the point is that it mostly is wrongful, or even that it in fact always is wrongful, in that it mostly is, or even in fact always is, unnecessary for the accomplishing of any morally acceptable wartime purpose—which leaves open that it is, or anyway might be, permissible when it is necessary for the accomplishing of a morally acceptable wartime purpose.<sup>11</sup> If that is correct, however, then terror bombing

11. This is the view taken by Michael Walzer in *Just and Unjust Wars* (New York: Basic Books, 1977), chap. 16. It seems to me very attractive.



is not in moral principle different from strategic or tactical bombing, for they too are wrongful when unnecessary for the accomplishing of a morally acceptable wartime purpose.)

I will have to bypass as too hard the question how the fact of war affects questions about self-defense. Let us take ourselves to have before us only the question what it is and is not permissible to do by way of self-defense in private life, when the episode is not part of a continuing process whose stakes are higher. That means we here bypass the question what makes it permissible for the pilot to drop his bombs in Strategic Bomber, and impermissible (if it is) for him to do so in Terror Bomber.

## VII

I had asked at the beginning of Section V whether it is everywhere impermissible to kill a bystander in defense of one's life. We looked in Sections V and VI at a putative counterexample to the idea that it is, namely, Strategic Bomber, and I have suggested that the permissibility of proceeding in that case is due to the fact of the war between the Goods and the Bads. Is it everywhere in private life impermissible to kill a bystander in defense of one's life? In the cases we looked at in Sections I through IV, what is in question is one life for one: yours and that of a person Y whom you have to kill if you are to save your life. We will look briefly in Section X at cases in which several lives would be saved by the killing of Y; let us consider here only cases in which what is in question is one life for one. I suggest that in all such cases, it is impermissible for X to kill Y if Y is a bystander.

But what exactly are bystanders, and why should it be thought that they are morally protected in this way?

A person is a bystander relative to a particular situation. Suppose Y is in no way causally involved in X's being at risk of death. That seems intuitively to be a sufficient condition for Y's being a bystander to the situation that consists in X's being at risk of death.

Is this condition necessary as well as sufficient? Ordinary use of the word "bystander" gives no precise directive. Perhaps we would regard Y as a bystander to the situation that consists in X's being at risk of death even if Y is in *some* way causally involved in it, so long as Y's causal involvement in it is in one or another respect minimal. No matter for our purposes. Two things are enough for us. First, if Y is in no way causally

involved in the situation that consists in X's being at risk of death, then Y is clearly a bystander to it. And second, if Y is causally involved in it, and not minimally so—as, for example, when it is Y himself or herself who is about to kill X—then Y is clearly not a bystander to it.

Then it suggests itself that we can explain as follows why bystanders are morally protected in cases in which what is in question is one life for one. Other things being equal, every person Y has a right against X that X not kill Y. That is vague, for what are those other things that may or may not be equal? But we do know about some of those other things. Suppose that X is at risk of death. Suppose also that Y is clearly a bystander to X's being at risk of death. Then the fact that X can save himself only by killing Y does *not* make other things be unequal, and thus does *not* make Y lack the right.

Might something else make Y lack the right? By hypothesis, we are considering only cases in which what is in question is one life for one, so that there are no other people whose life or limb is relevant.

Might something override Y's right? The fact that X can save himself only by killing Y no more overrides Y's right than it makes Y lack the right. I will in fact ignore the possibility of overridings since in all of the cases we are considering there is no better reason to think Y's right is overridden than there is to think Y lacks the right.

It seems to me very plausible then to think that, given that Y is clearly a bystander to the situation that consists in X's being at risk of death, Y has a right that X not kill Y, and that *that* is why X may not kill Y.

Consider, for example, the No cases we looked at in Section IV. You are at risk of death in all of those cases, and can save your life only by killing a bystander. Given that they are bystanders, the fact that you can save your life only by killing them does not make other things be unequal: it does not make them lack rights against you that you not kill them. Nor does anything else make them lack these rights. Then we can surely say that

(1) In the circumstances, he has a right that you not kill him

is true of each of them. If so, then a fortiori you may not kill them. Indeed, if these ideas are correct, we have an explanation of why you may not.

It is certainly arguable that others besides bystanders are morally protected in this way. I leave that open.

Let us now look at the two drivers in Villainous Aggressor and Innocent Aggressor, and the fat man in Innocent Threat. *They* are clearly not bystanders, so there is not that reason to think (1) true of each of them.

I have left open that others besides bystanders are morally protected in the way we are looking at, so perhaps there is some other reason to think (1) true of each of them?

Quite to the contrary, there is reason to think (1) false of them. As I said, other things being equal, every person Y has a right against X that X not kill Y. So in particular, other things being equal, *you* have rights against each of them that they not kill you. Suppose, in fact, that

(2) In the circumstances, you have a right that he not kill you  
is true of each of them. If so, then surely

(3) If he kills you, he will violate your right that he not kill you  
is also true of each of them. But given that

(4) If you do not kill him, he will kill you  
is also true of each of them, it surely follows that (1) is *not* true of any of them. That is, it follows that they lack rights that you not kill them. A fortiori, you may kill them. Indeed, if these ideas are correct, then we have an explanation of why you may.

I spelled all this out in detail in order that it be clear exactly where the shoe pinches.

The shoe fits well in Villainous Aggressor: we have no trouble regarding (2) and therefore (3) as true of the villainous driver, and we have no trouble regarding him as therefore lacking a right that you not kill him—he has, as we say, forfeited that right. But what of the driver of Innocent Aggressor? *He* is without fault, and we are accustomed to thinking of violation of rights as requiring fault in the violator.

But I suggest that the shoe stretches a bit with further wear, for there are good reasons to think we should not take fault to be required for a violation of a right. To begin with, the villainous driver in Villainous Aggressor has no right to kill you, and surely it is also true of the fault-free driver in Innocent Aggressor that *he* has no right to kill you. In Hohfeldian terms, neither of the two drivers has a privilege of killing you. For them to lack the privilege of killing you, however, is for you to have rights (Hohfeldian claims) that they not do so, rights they will infringe if they

succeed in killing you. So it really does seem right to think that (2) is true of the fault-free driver as well as of the villainous driver.

What of (3)? Well, given that (2) is true of the fault-free driver, you have a right that he not kill you. Can it be thought that morality all the same permits him to kill you? Can it be thought morally permissible for him to do so? No doubt the fault-free driver—unlike the villainous driver—will have an excuse, a very good excuse, a completely exculpatory excuse, if he runs you down; but these considerations suggest that he has no more than an excuse. If so, then he really does, however faultlessly, violate a right of yours if he kills you.

Can we conclude that (1) is therefore not true of the fault-free driver? Perhaps we will feel reluctant to say that the fault-free driver *forfeits* his right by virtue of being about to violate your right. Some people, I think, take it that forfeiting a right by definition requires fault. No matter. What is in question is not whether the innocent aggressor forfeits his right but whether he lacks it. And once we agree that he is about to violate your right—and that you can prevent this only by killing him—it seems right to conclude that he no longer has a right that you not kill him.

Still, there are bound to be some people who would resist these ideas: “No violation of a right without fault,” they would say. In particular, since the driver in *Innocent Aggressor* is free of fault, he will not violate a right of yours if you do not stop him, and he therefore does not cease to possess a right by virtue of doing what he does. But do these people then think it all the same permissible for you to kill the fault-free driver? How *could* they think this? I said in Section II that some people would insist that self-defense is merely excusable in *Innocent Aggressor*, and that I have a hypothesis as to what is at work in anyone of whom this is true. The hypothesis is precisely that they think the fault-free driver will violate no right of yours if you do not stop him, and therefore does not cease to possess a right by virtue of doing what he does. It would certainly be no wonder if a person who thought this also thought it merely excusable in you to blow up the truck in *Innocent Aggressor*. Moved by the considerations that suggest that fault is not required for violation of a right, I think it permissible for you to proceed. Others, as I say, may insist that fault is required. It constitutes confirmation of the account of self-defense I offered if they conclude that self-defense is not permissible, and is merely excusable, in *Innocent Aggressor*.

Does the shoe stretch far enough to accommodate *Innocent Threat* as

well? The fat man in that case is not only not at fault, he *does* nothing at all—he merely falls toward you. How can anyone say that (3) is true of him? I think there are good reasons in general to say that agency is no more required for violating a right than fault is, and good reasons therefore to say that (3) is true of the fat man, and therefore that (1) is not true of him.<sup>12</sup> No matter, if those who say “No violation of rights without agency” *also* say it is not permissible, but merely excusable, for you to proceed in Innocent Threat. I said in Section III that some people would insist that self-defense is merely excusable in Innocent Threat, and that I have a hypothesis as to what is at work in anyone of whom this is true. The hypothesis is precisely that they think the fat man will violate no right of yours if you do not stop him, and therefore does not cease to possess a right by virtue of falling toward you. It would certainly be no wonder if a person who thought this also thought it merely excusable in you to shift your awning. Moved by the considerations that suggest that agency is not required for violation of a right, I think it permissible for you to proceed. It constitutes confirmation of the account of self-defense I offered if those who think agency is required for violation of rights draw the conclusion that self-defense is not permissible, and is merely excusable, in Innocent Threat.

In short, I suggest that what makes it permissible for you to kill the two drivers and the fat man is the fact that they will otherwise violate your rights that they not kill you, and therefore lack rights that you not kill them.<sup>13</sup> Some people may say there is no such fact in the case of

12. Frances Kamm draws attention to the fact that if the falling fat man could alter his own direction of fall, then morality would call for his doing so, even at some considerable cost to himself—he is required to absorb a larger cost to prevent his landing on you than would be required of a third-party bystander who could prevent the fat man’s landing on you. (But she expresses doubt as to how large a cost: minor injury is one thing, death arguably quite another.) See her “The Insanity Defense and Innocent Threats,” *Criminal Justice Ethics* 6 (1987): 61–76, esp. pp. 63–67. The moral phenomenon she draws attention to is easily explained if we suppose that the fat man will violate a right of yours if he lands on you. (Kamm herself uses the terminology of “duty,” and may well have in mind the correlative of a Hohfeldian claim.)

13. I stress that what makes it permissible for you to kill them is the fact that they will otherwise violate your rights *that they not kill you*. That is a very stringent right. So is the right violated in rape, for example. But it would not be permissible for you to kill a man just on the ground that he will otherwise violate your right that he not take your wallet or your hat. Some of the lawyers whose views are analyzed by Fletcher (in “Proportionality and the Psychotic Aggressor”) think an appeal to rights violation is required to justify self-defense, and then seem to have had trouble understanding how there can be a ‘proportion-

either the fault-free driver or the fat man; some (probably more) people may say there is no such fact at least in the case of the fat man. All is well for the account of self-defense I offered if the first group are also content to say that it is impermissible (though excusable) for you to proceed in Innocent Aggressor and Innocent Threat, and the second group to say this at least of your proceeding in Innocent Threat.

## VIII

I hope it will have been noticed that the account I just offered of why you may proceed in Villainous Aggressor, Innocent Aggressor, and Innocent Threat is not the same as the account we had reached at the end of Section III. We were there considering the possibility of explaining why you may proceed in those cases by appeal to the premise “They will otherwise kill you.” The account I just offered includes that premise, but adds something, namely, that if they kill you they will violate your right that they not do so. That is a significant addition.

And it—or something that does a comparable job—is necessary, for it just is not sufficient to justify your killing a person that that person will otherwise kill you.<sup>14</sup>

---

ality’ requirement on self-defense; but it is curious that they should have had this trouble, since a proportionality requirement flows very naturally out of the fact that some rights are more stringent than others, a fact that any plausible theory of rights must accommodate. (Perhaps their having this trouble issues from their assuming that all rights are ‘absolute’, an assumption that would thereby be shown to have had bad effects in law as well as in philosophy.)

14. G.E.M. Anscombe believes that one may not directly kill the innocent, even in self-defense, but that the non-innocent are another matter. In *Mr. Truman’s Degree*, she defines “innocent” as “non-harming”; see note 1. That might incline us to think she believes “They will otherwise kill you” to be sufficient to justify your killing them. But in a later paper she defines “innocent” as follows: “What is required, for the people attacked to be non-innocent in the relevant sense, is that they should themselves be engaged in an objectively unjust proceeding which the attacker has the right to make his concern; or—the commonest case—should be unjustly attacking him. Then he can attack them with a view to stopping them. . . .” See G.E.M. Anscombe, “War and Murder,” reprinted in her *Collected Philosophical Papers* 3:51–61. If it is this latter that she all along had in mind as the non-innocence required in Y if X is to kill Y in defense of X’s life, then she at no time believed “They will otherwise kill you” sufficient to justify your killing them: she all along believed that justifying that conclusion requires something more, here expressed as the requirement that they be “engaged in an objectively unjust proceeding.” One way of interpreting that phrase is as requiring what I have offered, namely, that they be about to violate your rights; but however it is construed, the shoe will pinch in the same places—if it is to justify

Perhaps this will come out most vividly if we consider fighting back, a topic that any account of self-defense must in any case address itself to. Consider Villainous Aggressor, for example. You have an antitank gun, and it is permissible for you to blow up the truck. What if the villainous driver has his own antitank gun, and can use it on you so that you cannot use your antitank gun on him? Is it permissible for him to do this? I should think it obviously impermissible for him to do this, even if it is the case (since he will in fact continue to drive toward you in his truck) that you will shoot him unless he shoots you.

Why may the villainous driver not fight back in self-defense? We have an answer if we accept the account of self-defense that I offered in the preceding section. Suppose you are Alice, and that he shoots, and later says, "It was a matter of self-defense, your honor; after all, Alice was going to kill me unless I killed her." Self-defense? How so? Suppose he says, "Look, your honor, it might have been thought that

(1') In the circumstances, Alice had a right that I not kill her  
is true. But it is not true. For

(2') In the circumstances, I had a right that Alice not kill me,  
and therefore

(3') If Alice killed me she would violate my right that she not kill me  
are true. But

(4') If I did not kill her, she was going to kill me

is also true. It follows that (1') is not true. It follows that it was permissible for me to kill her." This is a bad joke. (2') is false, for he had ceased to have the right by virtue of driving toward you in the truck.

Is fighting back permissible in Innocent Aggressor and Innocent Threat? I think not, but that is because I think the same holds of the fault-free driver and the fat man as holds of the villainous driver, namely, that they too lack rights that you not kill them. Others may disagree. It would be a further confirmation of the account of self-defense I offered if those who think the fault-free driver and the fat man do have rights that you not kill them also think it permissible for them to fight back.

---

your proceeding in Innocent Aggressor and Innocent Threat as well as in Villainous Aggressor. (I suspect that the role of that "objectively" is precisely to stretch the shoe.)

In short, I suggest that the permissibility of X's killing Y in self-defense goes hand in hand with the impermissibility of Y's fighting back, and that both phenomena have a common source.

I am of course assuming it to be obvious that the villainous driver may not fight back. I suspect that some would take a Hobbesian line here: on their view, all bets are off when a person will otherwise kill you. Thus they would say that the premise "Y will otherwise kill X" really is sufficient for the conclusion that X may kill Y, so that even the villainous driver may fight back. But why stop there? Why not go on to say more generally that all bets are off when you will otherwise *die*, and thus that the premise "X will otherwise die" is sufficient for the conclusion that X may kill Y? That, alas, yields that you may eat the baby in Starvation. It is hard to see what rationale could be offered for stopping this slide into moral absurdity, and that should make us suspicious of what inclined us to want to start it. In any case, I am assuming that we do not want to start it.

## IX

All of the cases we have been looking at are cases in which the situation you find yourself in was ultimately caused by a villain. All of them (except, I suppose, Villainous Aggressor) have variants in which no villain is involved at all. Thus, for example, we might imagine a case in which the fat man in Innocent Threat was caused to fall toward you, not by a villain, but instead by a sudden strong gust of wind. Does the presence or absence of a villain make a difference? I should think not. That it is a villain as opposed to nature (or as opposed to someone without fault) who made it be the case that you have to kill a person to save your life does not by itself work in support of its being the case that you may proceed—or in support of its being the case that you may not.

But there are two other kinds of variants on those cases that are worth stopping over.

In the first place there are what might be called Third-Party cases, in which what is in question is not self-defense, but other-defense: cases in which you cannot save yourself but some third party can. So, for example, let us imagine that you have no antitank gun in Villainous Aggressor, but a third party does; then what we have is a Third-Party case, and our question is whether the third party may use *his* antitank gun to save



*your* life, and if so, why. Similar variants are available for all of the cases we have been looking at.

Self- and other-defense are not exactly two sides of one coin, but they are nevertheless close to it. I say they are not exactly two sides of one coin in light of considerations of autonomy. Suppose that X can defend himself or herself against an attack by Y and that a third party Z can also defend X against the attack by Y. In some such cases, Z ought to refrain from proceeding in that defending X would constitute “barging in,” meddling, interfering. Indeed, Z’s defending X might constitute “barging in” even when X cannot defend himself against the attack by Y—we do think that people should be left to fight their own battles, even at the cost of losing some. (Some, but decidedly not all. We draw a line well before a potential victim’s life is at stake.)

Considerations of autonomy apart, however, I think it very plausible to suppose that the permissibility of X’s killing Y in self-defense goes hand in hand with the permissibility of Z’s killing Y in defense of X, and that both phenomena have a common source. So all three phenomena—the permissibility of X’s killing Y in self-defense, the impermissibility of Y’s fighting back, and the permissibility of Z’s killing Y in defense of X—go hand in hand and have a common source.

And I take it to be a further confirmation of the account of self-defense I offered if anyone who thinks the fault-free driver and the fat man have rights that you not kill them also thinks they have rights that a third party not kill them, and thus that it is impermissible for the third party to intervene in your defense.

Some people may say, “But the driver in Innocent Aggressor and the fat man in Innocent Threat are without fault!” Of course they are without fault. But why should that bar third-party action if it does not bar first-party action?

Some people may ask, “Who is the third party to choose who shall live as between two who are equally without fault?” You and the driver are both without fault in Innocent Aggressor; you and the fat man are both without fault in Innocent Threat. But why should that bar third-party choice if it does not bar first-party choice?

In addition, it should be remembered that if the driver in Innocent Aggressor and the fat man in Innocent Threat are about to violate your rights, then they are not merely people who are without fault. Not that the third party must intervene: I should think he may prefer to do noth-

ing at all. (Just as you may prefer to do nothing at all.) If, on the other hand, they are not about to violate your rights, then it is impermissible for the third party to intervene—just as it is impermissible for you to do so.

The account of self-defense that I gave in Section VII may of course be incorrect. But if the facts in *Innocent Aggressor* and *Innocent Threat* did not in *some* way make it permissible for the third party to intervene in your favor, how *could* they make it permissible for you to intervene in your own favor?

One kind of answer that has been offered is this: what makes it permissible for you to kill the driver in *Innocent Aggressor* and the fat man in *Innocent Threat* is the fact that you have an “agent-relative permission” to do so, a permission that the third party lacks. The idea is that X (and only X) is permitted to kill Y when Y will otherwise kill X, since it is X (and only X) whom Y will otherwise kill.<sup>15</sup> But that cannot be right.

In the first place, not just anything goes in self-defense: we have no across-the-board agent-relative permission to kill others to save our lives. (You certainly may not eat the baby in *Starvation*.) Could it be said that we have an agent-relative permission to kill another to save our lives but that the permission is limited to cases in which the other will otherwise kill us? What is supposed to be the rationale for this limit? If agent-relativity is what is crucial (Me! Me! I’m the one at risk!), why *isn’t* the permission an across-the-board permission? (Compare my comments in the preceding section about the Hobbesian idea that all bets are off when Y will kill X unless X kills Y.)

Moreover, if the agent-relative permission is limited in this way, why *isn’t* there a comparable limited third-party-centered permission? Is it just to be taken, flatly, as a datum, that while I have a permission to kill a person who will otherwise kill me, a third party lacks a permission to kill a person who will otherwise kill me?

Most important, it just is not acceptable to say that a person may do to save his own life something that it would not be permissible for him or her to do to save the life of another—considerations of the other’s autonomy apart—for whatever I may do to save my own life I may surely do to save the life of someone I love.

15. Nancy Davis offers this kind of answer in regard to a similar case in “Abortion and Self-Defense,” *Philosophy & Public Affairs* 13, no. 3 (Summer 1984): 175–207. As her title indicates, she brings out the bearing of these issues on the problem of abortion.

Well, on some views of the agent-relative permission it extends to defending those I love as well as to defending myself. (After all, those I love are those *I* love, so that it's still Me! Me! that is crucial.) But do we really wish to have it that a man's being rather a bore is going to make him later be undefendable in circumstances in which he would have been defendable had he only been charming enough to be loved?

No doubt X may have an excuse for doing something to save his or her own life that it was wrongful in X to do to save it, when a third party who saved X's life in those circumstances would have acted not merely wrongly but also inexcusably. (Just as X may have an excuse for doing something wrongful to save his or her child's life, when a third party who saved X's child's life in those circumstances would have acted not merely wrongly but also inexcusably.) So far as I can see, this is the only *fact* that lies behind appeals to agent-relative permission here. And it does not come to much.<sup>16</sup> For having an excuse for doing something wrongful is just that: doing something wrongful, though having an excuse for doing it. But our question was whether it is permissible (and not merely excusable) for you to proceed in these cases.

In short, it is not because of the personal fact of your special relation to those who will kill you in those cases, it is not because they will otherwise kill *you*, that you may proceed; it is because of the entirely impersonal fact that they will *otherwise violate your rights that they not kill you* that you may proceed. But that impersonal fact may be acted on by a third party as well as by you.

## X

A second interesting kind of variant on these cases is what we get when we imagine that what is in question is not self-defense but other-defense, and moreover, not merely other-defense but other-defense of more than one. So, for example, let us imagine that the villain in Villainous Aggres-

16. I should perhaps make clear that I am not saying that appeals to agent-relative permissions are unsuccessful everywhere in moral theory. For example, they certainly do seem to have a role to play in cases that involve only the distribution of benefits, i.e., in cases of the kind discussed by John M. Taurek in "Should the Numbers Count?" *Philosophy & Public Affairs* 6, no. 4 (Summer 1977): 293-316.

sor is driving toward five, not one, and that none of the five has an anti-tank gun. Is it permissible for a third party to use his antitank gun to save the five? Yes. If it is permissible for a third party to use his antitank gun to save only one, it is surely permissible for him to do so to save five. Similarly for similar variants on Innocent Aggressor and Innocent Threat. Let us call these Third-Party Different Number cases.

What is of interest here are the Third-Party Different Number variants on the No cases we surveyed in Section IV. In Trolley, a villain has started a trolley down a track toward you. You cannot stop the trolley, but you can deflect it. Unfortunately, the only path onto which you can deflect it will take it onto a bystander. It is impermissible for you to deflect the trolley. In a Third-Party variant on this case, you cannot do anything at all to save yourself, but a third party can save you—he can deflect the trolley onto the bystander. Given that it is impermissible for you to deflect the trolley in Trolley, the thesis of the preceding section tells us it is also impermissible for the third party to deflect the trolley in the Third-Party variant of Trolley. That consequence of the thesis is surely welcome. But now consider a Third-Party Different Number variant of Trolley in which the third party will save five if he deflects the trolley.<sup>17</sup> In *this* case it would be permissible for the third party to deflect the trolley: the third party's turning the trolley here would yield a net saving of four lives. Well, not just anything is permissible on the ground that it yields a net saving of four lives (supplying five who are starving with knives, forks, and the neighbors' baby obviously will not do), and what has come to be known as the Trolley Problem is the problem of explaining why it is permissible to do what would yield this net saving of lives—when it is.

I do not address that problem here, except to draw attention to the fact that a satisfactory solution to it must make clear why altering the numbers makes a difference. That is, it is *not* permissible for the third party to proceed in a variant on Trolley in which only one will be saved; it *is* permissible for the third party to proceed if five instead of one will be saved. The answer to the question why it is permissible to proceed when five will be saved had better be sensitive to the numbers: it had better

17. The case comes, of course, from Philippa Foot's "The Problem of Abortion and the Doctrine of the Double Effect," reprinted in her *Virtues and Vices* (Oxford: Basil Blackwell, 1978), pp. 19–32.

explain why it matters that five instead of one will be saved.<sup>18</sup> It is not enough to say *just*, “Well, five is four more than one”; we need an account of why adding the four makes a difference. Solutions to the Trolley Problem that focus on the means the agent uses to save those he saves do not, and I think cannot, answer to this need.

18. I mention with some diffidence that the solution to this problem offered in my *The Realm of Rights* (Cambridge, Mass.: Harvard University Press, 1990) is sensitive to the numbers, for what that solution focuses on is the increased probability of advantage secured for all six by a decision made in advance to turn the trolley when it later comes to threaten five and can be turned onto a sixth.

For what it is worth, however, I should mention that my discussion of self-defense here departs from that of chap. 14 of that book: I now think I was mistaken about some of the cases discussed there.