The Trolley Problem

Author(s): Judith Jarvis Thomson

Source: *The Yale Law Journal*, May, 1985, Vol. 94, No. 6 (May, 1985), pp. 1395-1415

Published by: The Yale Law Journal Company, Inc.

Stable URL: http://www.jstor.com/stable/796133

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms



is collaborating with JSTOR to digitize, preserve and extend access to *The Yale Law Journal*

This content downloaded from 12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

# Comments

# The Trolley Problem

**Judith Jarvis Thomson†**

## I.

Some years ago, Philippa Foot drew attention to an extraordinarily interesting problem.[1] Suppose you are the driver of a trolley. The trolley rounds a bend, and there come into view ahead five track workmen, who have been repairing the track. The track goes through a bit of a valley at that point, and the sides are steep, so you must stop the trolley if you are to avoid running the five men down. You step on the brakes, but alas they don't work. Now you suddenly see a spur of track leading off to the right. You can turn the trolley onto it, and thus save the five men on the straight track ahead. Unfortunately, Mrs. Foot has arranged that there is one track workman on that spur of track. He can no more get off the track in time than the five can, so you will kill him if you turn the trolley onto him. Is it morally permissible for you to turn the trolley?

Everybody to whom I have put this hypothetical case says, Yes, it is.[2] Some people say something stronger than that it is morally *permissible* for you to turn the trolley: They say that morally speaking, you *must* turn it—that morality requires you to do so. Others do not agree that morality

---

† Professor of Philosophy, Massachusetts Institute of Technology. B.A., Barnard College 1950, Cambridge University 1952; M.A., Cambridge University 1956; Ph.D., Columbia University 1959. Many people have given me helpful criticism of this essay's many successive reincarnations over the years; I cannot list them all—for want of space, not of gratitude. Most recently, it benefited from criticism by the members of the Yale Law School Civil Liability Workshop and the Legal Theory Workshop, Faculty of Law, University of Toronto.

1. *See* P. FOOT, *The Problem of Abortion and the Doctrine of the Double Effect*, in VIRTUES AND VICES AND OTHER ESSAYS IN MORAL PHILOSOPHY 19 (1978).

2. I think it possible (though by no means certain) that John Taurek would say No, it is not permissible to (all simply) turn the trolley; what you ought to do is flip a coin. *See* Taurek, *Should the Numbers Count?*, 6 PHIL. & PUB. AFF. 293 (1977). (But he is there concerned with a different kind of case, namely that in which what is in question is not whether we may do what harms one to avoid harming five, but whether we may or ought to choose to save five in preference to saving one.) For criticism of Taurek's article, see Parfit, *Innumerate Ethics*, 7 PHIL. & PUB. AFF. 285 (1978).



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

requires you to turn the trolley, and even feel a certain discomfort at the idea of turning it. But everybody says that it is true, at a minimum, that you *may* turn it—that it would not be morally wrong in you to do so.

Now consider a second hypothetical case. This time you are to imagine yourself to be a surgeon, a truly great surgeon. Among other things you do, you transplant organs, and you are such a great surgeon that the organs you transplant always take. At the moment you have five patients who need organs. Two need one lung each, two need a kidney each, and the fifth needs a heart. If they do not get those organs today, they will all die; if you find organs for them today, you can transplant the organs and they will all live. But where to find the lungs, the kidneys, and the heart? The time is almost up when a report is brought to you that a young man who has just come into your clinic for his yearly check-up has exactly the right blood-type, and is in excellent health. Lo, you have a possible donor. All you need do is cut him up and distribute *his* parts among the five who need them. You ask, but he says, "Sorry. I deeply sympathize, but no." Would it be morally permissible for you to operate anyway? Everybody to whom I have put this second hypothetical case says, No, it would not be morally permissible for you to proceed.

Here then is Mrs. Foot's problem: *Why* is it that the trolley driver may turn his trolley, though the surgeon may not remove the young man's lungs, kidneys, and heart?[3] In both cases, one will die if the agent acts, but five will live who would otherwise die—a net saving of four lives. What difference in the other facts of these cases explains the moral difference between them? I fancy that the theorists of tort and criminal law will find this problem as interesting as the moral theorist does.

## II.

Mrs. Foot's own solution to the problem she drew attention to is simple, straightforward, and very attractive. She would say: Look, the surgeon's choice is between operating, in which case he kills one, and not operating, in which case he lets five die; and killing is surely worse than letting die[4]—indeed, so much worse that we can even say

(I) Killing one is worse than letting five die.

---

3. I doubt that anyone would say, with any hope of getting agreement from others, that the surgeon ought to flip a coin. So even if you think that the trolley driver ought to flip a coin, there would remain, for you, an analogue of Mrs. Foot's problem, namely: Why ought the trolley driver flip a coin, whereas the surgeon may not?

4. Mrs. Foot speaks more generally of causing injury and failing to provide aid; and her reason for thinking that the former is worse than the latter is that the negative duty to refrain from causing injury is stricter than the positive duty to provide aid. *See* P. FOOT, *supra* note 1, at 27–29.



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

So the surgeon must refrain from operating. By contrast, the trolley driver's choice is between turning the trolley, in which case he kills one, and not turning the trolley, in which case he does not *let five die*, he positively *kills* them. Now surely we can say

(II) Killing five is worse than killing one.

But then that is why the trolley driver may turn his trolley: He would be doing what is worse if he fails to turn it, since if he fails to turn it he kills five.

I do think that that is an attractive account of the matter. It seems to me that if the surgeon fails to operate, he does not kill his five patients who need parts; he merely lets them die. By contrast, if the driver fails to turn his trolley, he does not merely let the five track workmen die; he drives his trolley into them, and thereby kills them.

But there is good reason to think that this problem is not so easily solved as that.

Let us begin by looking at a case that is in some ways like Mrs. Foot's story of the trolley driver. I will call her case *Trolley Driver*; let us now consider a case I will call *Bystander at the Switch*. In that case you have been strolling by the trolley track, and you can see the situation at a glance: The driver saw the five on the track ahead, he stamped on the brakes, the brakes failed, so he fainted. What to do? Well, here is the switch, which you can throw, thereby turning the trolley yourself. Of course you will kill one if you do. But I should think you may turn it all the same.[5]

Some people may feel a difference between these two cases. In the first place, the trolley driver is, after all, captain of the trolley. He is charged by the trolley company with responsibility for the safety of his passengers and anyone else who might be harmed by the trolley he drives. The bystander at the switch, on the other hand, is a private person who just happens to be there.

Second, the driver would be driving a trolley into the five if he does not turn it, and the bystander would not—the bystander will do the five no harm at all if he does not throw the switch.

I think it right to feel these differences between the cases.

Nevertheless, my own feeling is that an ordinary person, a mere bystander, may intervene in such a case. If you see something, a trolley, a boulder, an avalanche, heading towards five, and you can deflect it onto

5. A similar case (intended to make a point similar to the one that I shall be making) is discussed in Davis, *The Priority of Avoiding Harm*, in KILLING AND LETTING DIE 172, 194–95 (B. Steinbock ed. 1980).



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

one, it really does seem that—other things being equal—it would be permissible for you to *take* charge, *take* responsibility, and deflect the thing, whoever you may be. Of course you run a moral risk if you do, for it might be that, unbeknownst to you, other things are not equal. It might be, that is, that there is some relevant difference between the five on the one hand, and the one on the other, which would make it morally preferable that the five be hit by the trolley than that the one be hit by it. That would be so if, for example, the five are not track workmen at all, but Mafia members in workmen's clothing, and they have tied the one workman to the right-hand track in the hope that you would turn the trolley onto him. I won't canvass all the many kinds of possibilities, for in fact the moral risk is the same whether you are the trolley driver, or a bystander at the switch.

Moreover, second, we might well wish to ask ourselves what exactly is the difference between what the driver would be doing if he failed to turn the trolley and what the bystander would be doing if he failed to throw the switch. As I said, the driver would be driving a trolley into the five; but what exactly would his driving the trolley into the five consist in? Why, just sitting there, doing nothing! If the driver does just sit there, doing nothing, then that will have been how come he drove his trolley into the five.

I do not mean to make much of that fact about what the driver's driving his trolley into the five would consist in, for it seems to me to be right to say that if he does not turn the trolley, he does drive his trolley into them, and does thereby kill them. (Though this does seem to me to be right, it is not easy to say exactly what makes it so.) By contrast, if the bystander does not throw the switch, he drives no trolley into anybody, and he kills nobody.

But as I said, my own feeling is that the bystander *may* intervene. Perhaps it will seem to some even less clear that morality requires him to turn the trolley than that morality requires the driver to turn the trolley; perhaps some will feel even more discomfort at the idea of the bystander's turning the trolley than at the idea of the driver's turning the trolley. All the same, I shall take it that he *may*.

If he may, there is serious trouble for Mrs. Foot's thesis (I). It is plain that if the bystander throws the switch, he causes the trolley to hit the one, and thus he kills the one. It is equally plain that if the bystander does not throw the switch, he does not cause the trolley to hit the five, he does not kill the five, he merely fails to save them—he lets them die. His choice therefore is between throwing the switch, in which case he kills one, and not throwing the switch, in which case he lets five die. If thesis (I) were



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

true, it would follow that the bystander may not throw the switch, and that I am taking to be false.

## III.

I have been arguing that

(I) Killing one is worse than letting five die

is false, and a fortiori that it cannot be appealed to to explain why the surgeon may not operate in the case I shall call *Transplant.*

I think it pays to take note of something interesting which comes out when we pay close attention to

(II) Killing five is worse than killing one.

For let us ask ourselves how we would feel about *Transplant* if we made a certain addition to it. In telling you that story, I did not tell you why the surgeon's patients are in need of parts. Let us imagine that the history of their ailments is as follows. The surgeon was badly overworked last fall—some of his assistants in the clinic were out sick, and the surgeon had to take over their duties dispensing drugs. While feeling particularly tired one day, he became careless, and made the terrible mistake of dispensing chemical X to five of the day's patients. Now chemical X works differently in different people. In some it causes lung failure, in others kidney failure, in others heart failure. So these five patients who now need parts need them because of the surgeon's carelessness. Indeed, if he does not get them the parts they need, so that they die, he will have killed them. Does that make a moral difference? That is, does the fact that he will have killed the five if he does nothing make it permissible for him to cut the young man up and distribute his parts to the five who need them?

We could imagine it to have been worse. Suppose what had happened was this: The surgeon was badly overextended last fall, he had known he was named a beneficiary in his five patients' wills, and it swept over him one day to give them chemical X to kill them. Now he repents, and would save them if he could. If he does not save them, he will positively have murdered them. Does *that* fact make it permissible for him to cut the young man up and distribute his parts to the five who need them?

I should think plainly not. The surgeon must not operate on the young man. If he can find no other way of saving his five patients, he will *now* have to let them die—despite the fact that if he now lets them die, he will have killed them.



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

We tend to forget that some killings themselves include lettings die, and do include them where the act by which the agent kills takes time to cause death—time in which the agent can intervene but does not.

In face of these possibilities, the question arises what we should think of thesis (II), since it *looks* as if it tells us that the surgeon ought to operate, and thus that he may permissibly do so, since if he operates he kills only one instead of five.

There are two ways in which we can go here. First, we can say: (II) does tell us that the surgeon ought to operate, and that shows it is false. Second, we can say: (II) does not tell us that the surgeon ought to operate, and it is true.

For my own part, I prefer the second. If Alfred kills five and Bert kills only one, then questions of motive apart, and other things being equal, what Alfred did *is* worse than what Bert did. If the surgeon does not operate, so that he kills five, then it will later be true that he did something worse than he would have done if he had operated, killing only one—especially if his killing of the five was murder, committed out of a desire for money, and his killing of the one would have been, though misguided and wrongful, nevertheless a well-intentioned effort to save five lives. Taking this line would, of course, require saying that assessments of which acts are worse than which other acts do not by themselves settle the question what it is permissible for an agent to do.

But it might be said that we ought to by-pass (II), for perhaps what Mrs. Foot would have offered us as an explanation of why the driver may turn the trolley in *Trolley Driver* is not (II) itself, but something more complex, such as

> (II′) If a person is faced with a choice between doing something *here and now* to five, by the doing of which he will kill them, and doing something else *here and now* to one, by the doing of which he will kill only the one, then (other things being equal) he ought to choose the second alternative rather than the first.

We may presumably take (II′) to tell us that the driver ought to, and hence permissibly may, turn the trolley in *Trolley Driver*, for we may presumably view the driver as confronted with a choice between here and now driving his trolley into five, and here and now driving his trolley into one. And at the same time, (II′) tells us nothing at all about what the surgeon ought to do in *Transplant*, for he is not confronted with such a choice. If the surgeon operates, he does do something by the doing of which he will kill only one; but if the surgeon does not operate, he does not do something by the doing of which he kills five; he merely fails to do



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

something by the doing of which he would make it be the case that he has not killed five.

I have no objection to this shift in attention from (II) to (II′). But we should not overlook an interesting question that lurks here. As it might be put: *Why* should the present tense matter so much? Why should a person prefer killing one to killing five if the alternatives are wholly in front of him, but not (or anyway, not in every case) where one of them is partly behind him? I shall come back to this question briefly later.

Meanwhile, however, even if (II′) can be appealed to in order to explain why the trolley driver may turn his trolley, that would leave it entirely open why the bystander at the switch may turn *his* trolley. For he does not drive a trolley into each of five if he refrains from turning the trolley; he merely lets the trolley drive into each of them.

So I suggest we set *Trolley Driver* aside for the time being. What I shall be concerned with is a first cousin of Mrs. Foot's problem, viz.: Why is it that the bystander may turn his trolley, though the surgeon may not remove the young man's lungs, kidneys, and heart? Since *I* find it particularly puzzling that the bystander may turn his trolley, I am inclined to call this The Trolley Problem. Those who find it particularly puzzling that the surgeon may not operate are cordially invited to call it The Transplant Problem instead.

## IV.

It should be clear, I think, that "kill" and "let die" are too blunt to be useful tools for the solving of this problem. We ought to be looking within killings and savings for the ways in which the agents would be carrying them out.

It would be no surprise, I think, if a Kantian idea occurred to us at this point. Kant said: "Act so that you treat humanity, whether in your own person or in that of another, always as an end and never as a means only." It is striking, after all, that the surgeon who proceeds in *Transplant* treats the young man he cuts up "as a means only": He literally uses the young man's body to save his five, and does so without the young man's consent. And perhaps we may say that the agent in *Bystander at the Switch* does not use his victim to save his five, or (more generally) treat his victim as a means only, and that that is why he (unlike the surgeon) may proceed.

But what exactly is it to treat a person as a means only, or to use a person? And why exactly is it wrong to do this? These questions do not have obvious answers.[6]

6. For a sensitive discussion of some of the difficulties, see Davis, *Using Persons and Common*

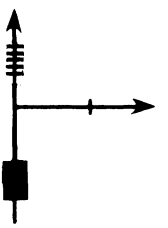


This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

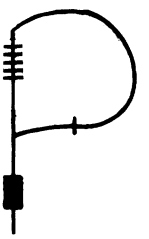
Suppose an agent is confronted with a choice between doing nothing, in which case five die, or engaging in a certain course of action, in which case the five live, but one dies. Then perhaps we can say: If the agent chooses to engage in the course of action, then he uses the one to save the five only if, had the one gone out of existence just before the agent started, the agent would have been unable to save the five. That is true of the surgeon in *Transplant.* He needs the young man if he is to save his five; if the young man goes wholly out of existence just before the surgeon starts to operate, then the surgeon cannot save his five. By contrast, the agent in *Bystander at the Switch* does not need the one track workman on the right-hand track if he is to save his five; if the one track workman goes wholly out of existence before the bystander starts to turn the trolley, then the bystander *can* all the same save his five. So here anyway is a striking difference between the cases.

It does seem to me right to think that solving this problem requires attending to the means by which the agent would be saving his five if he proceeded. But I am inclined to think that this is an overly simple way of taking account of the agent's means.

One reason for thinking so[7] comes out as follows. You have been thinking of the tracks in *Bystander at the Switch* as not merely diverging, but continuing to diverge, as in the following picture: pick up figure 1



Consider now what I shall call "the loop variant" on this case, in which the tracks do not continue to diverge—they circle back, as in the following picture:



*Sense*, 94 ETHICS 387 (1984). Among other things, she argues (I think rightly) that the Kantian idea is not to be identified with the common sense concept of "using a person." *Id.* at 402.

7. For a second reason to think so, see *infra* note 13.



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

Let us now imagine that the five on the straight track are thin, but thick enough so that although all five will be killed if the trolley goes straight, the bodies of the five will stop it, and it will therefore not reach the one. On the other hand, the one on the right-hand track is fat, so fat that his body will by itself stop the trolley, and the trolley will therefore not reach the five. May the agent turn the trolley? Some people feel more discomfort at the idea of turning the trolley in the loop variant than in the original *Bystander at the Switch*. But we cannot really suppose that the presence or absence of that extra bit of track makes a major moral difference as to what an agent may do in these cases, and it really does seem right to think (despite the discomfort) that the agent may proceed.

On the other hand, we should notice that the agent here needs the one (fat) track workman on the right-hand track if he is to save his five. If the one goes wholly out of existence just before the agent starts to turn the trolley, then the agent cannot save his five[8]—just as the surgeon in *Transplant* cannot save his five if the young man goes wholly out of existence just before the surgeon starts to operate.

Indeed, I should think that there is no plausible account of what is involved in, or what is necessary for, the application of the notions "treating a person as a means only," or "using one to save five," under which the surgeon would be doing this whereas the agent in this variant of *Bystander at the Switch* would not be. If that is right, then appeals to these notions cannot do the work being required of them here.

## V.

Suppose the bystander at the switch proceeds: He throws the switch, thereby turning the trolley onto the right-hand track, thereby causing the one to be hit by the trolley, thereby killing him—but saving the five on the straight track. There are two facts about what he does which seem to me to explain the moral difference between what he does and what the agent in *Transplant* would be doing if *he* proceeded. In the first place, the bystander saves his five by making something that threatens them instead threaten one. Second, the bystander does not do that by means which themselves constitute an infringement of any right of the one's.

As is plain, then, my hypothesis as to the source of the moral difference between the cases makes appeal to the concept of a right. My own feeling

8. It is also true that if the five go wholly out of existence just before the agent starts to turn the trolley, then the one will die whatever the agent does. Should we say, then, that the agent uses one to save five if he acts, *and* uses five to save one if he does not act? No: What follows *and* is false. If the agent does not act, he uses nobody. (I doubt that it can even be said that if he does not act, he lets them *be used*. For what is the active for which this is passive? Who or what would be using them if he does not act?).



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

is that solving this problem requires making appeal to that concept—or to some other concept that does the same kind of work.[9] Indeed, I think it is one of the many reasons why this problem is of such interest to moral theory that it does force us to appeal to that concept; and by the same token, that we learn something from it about that concept.

Let us begin with an idea, held by many friends of rights, which Ronald Dworkin expressed crisply in a metaphor from bridge: Rights "trump" utilities.[10] That is, if one would infringe a right in or by acting, then it is not sufficient justification for acting that one would thereby maximize utility. It seems to me that something like this must be correct.

Consideration of this idea suggests the possibility of a very simple solution to the problem. That is, it might be said (i) The reason why the surgeon may not proceed in *Transplant* is that if he proceeds, he maximizes utility, for he brings about a net saving of four lives, but in so doing he would infringe a right of the young man's.

Which right? Well, we might say: The right the young man has against the surgeon that the surgeon not kill him—thus a right in the cluster of rights that the young man has in having a right to life.

Solving this problem requires being able to explain also why the bystander may proceed in *Bystander at the Switch*. So it might be said (ii) The reason why the bystander may proceed is that if he proceeds, he maximizes utility, for he brings about a net saving of four lives, and in so doing he does *not* infringe any right of the one track workman's.

But I see no way—certainly there is no easy way—of establishing that these ideas are true.

Is it clear that the bystander would infringe no right of the one track workman's if he turned the trolley? Suppose there weren't anybody on the straight track, and the bystander turned the trolley onto the right-hand track, thereby killing the one, but not saving anybody, since nobody was at risk, and thus nobody needed saving. Wouldn't that infringe a right of the one workman's, a right in the cluster of rights that he has in having a right to life?

So should we suppose that the fact that there are five track workmen on the straight track who are in need of saving makes the one lack that right—which he would have had if that had not been a fact?

But then why doesn't the fact that the surgeon has five patients who are in need of saving make the young man also lack that right?

I think some people would say there is good (excellent, conclusive) reason for thinking that the one track workman lacks the right (given there

9. I strongly suspect that giving an account of what makes it wrong to *use* a person, *see supra* text accompanying notes 6–8, would also require appeal to the concept of a right.

10. R. DWORKIN, TAKING RIGHTS SERIOUSLY ix (1977).



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

are five on the straight track) lying in the fact that (given there are five on the straight track) it is morally permissible to turn the trolley onto him. But if your reason for thinking the one lacks the right is that it is permissible to turn the trolley onto him, then you can hardly go on to explain its being permissible to turn the trolley onto him by appeal to the fact that he lacks the right. It pays to stress this point: If you want to say, as (ii) does, that the bystander may proceed because he maximizes utility and infringes no right, then you need an independent account of what makes it be the case that he infringes no right—independent, that is, of its being the case that he may proceed.

There is *some* room for maneuver here. Any plausible theory of rights must make room for the possibility of waiving a right, and within that category, for the possibility of failing to have a right by virtue of assumption of risk; and it might be argued that that is what is involved here, i.e., that track workmen know of the risks of the job, and consent to run them when signing on for it.

But that is not really an attractive way of dealing with this difficulty. Track workmen certainly do not explicitly consent to being run down with trolleys when doing so will save five who are on some other track—certainly they are not asked to consent to this at the time of signing on for the job. And I doubt that they consciously assume the risk of it at that or any other time. And in any case, what if the six people involved had not been track workmen? What if they had been young children? What if they had been people who had been shoved out of helicopters? Wouldn't it all the same be permissible to turn the trolley?

So it is not clear what (independent) reason could be given for thinking that the bystander will infringe no right of the one's if he throws the switch.

I think, moreover, that there is *some* reason to think that the bystander will infringe a right of the one if he throws the switch, even though it is permissible for him to do so. What I have in mind issues simply from the fact that if the bystander throws the switch, then he does what will kill the one. Suppose the bystander proceeds, and that the one is now dead. The bystander's motives were, of course, excellent—he acted with a view to saving five. But the one did not volunteer his life so that the five might live; the bystander volunteered it for him. The bystander made him pay with his life for the bystander's saving of the five. This consideration seems to me to lend some weight to the idea that the bystander did do him a wrong—a wrong it was morally permissible to do him, since five were saved, but a wrong *to him* all the same.

Consider again that lingering feeling of discomfort (which, as I said, some people do feel) about what the bystander does if he turns the trolley.



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

No doubt it is permissible to turn the trolley, but still . . . but still . . . . People who feel this discomfort also think that, although it is permissible to turn the trolley, it is not morally required to do so. My own view is that they are right to feel and think these things. We would be able to explain why this is so if we supposed that if the bystander turns the trolley, then he does do the one track workman a wrong—if we supposed, in particular, that he infringes a right of the one track workman's which is in that cluster of rights which the workman has in having a right to life.[11]

I do not for a moment take myself to have established that (ii) is false. I have wished only to draw attention to the difficulty that lies ahead of a person who thinks (ii) true, and also to suggest that there is some reason to think that the bystander would infringe a right of the one's if he proceeded, and thus some reason to think that (ii) is false. It can easily be seen that if there is some reason to think the bystander would infringe a right of the one's, then there is also some reason to think that (i) is false—since if the bystander does infringe a right of the one's if he proceeds, and may nevertheless proceed, then it cannot be the fact that the surgeon infringes a right of the young man's if *he* proceeds which makes it impermissible for *him* to do so.

Perhaps a friend of (i) and (ii) can establish that they are true. I propose that, just in case he can't, we do well to see if there isn't some other way of solving this problem than by appeal to them. In particular, I propose we grant that both the bystander and the surgeon would infringe a right of their ones, a right in the cluster of rights that the ones' have in having a right to life, and that we look for some *other* difference between the cases which could be appealed to to explain the moral difference between them.

Notice that accepting this proposal does not commit us to rejecting the idea expressed in that crisp metaphor of Dworkin's. We can still say that rights trump utilities—if we can find a further feature of what the bystander does if he turns the trolley (beyond the fact that he maximizes utility) which itself trumps the right, and thus makes it permissible to proceed.

## VI.

As I said, my own feeling is that the trolley problem can be solved only by appeal to the concept of a right—but not by appeal to it in as simple a way as that discussed in the preceding section. What we were attending to

11. Many of the examples discussed by Bernard Williams and Ruth Marcus plainly call out for this kind of treatment. *See* B. WILLIAMS, *Ethical Consistency*, in PROBLEMS OF THE SELF 166 (1973); Marcus, *Moral Dilemmas and Consistency*, 77 J. PHIL. 121 (1980).



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

in the preceding section was only the fact that the agents would be killing and saving if they proceeded; what we should be attending to is the means by which they would kill and save.[12] (It is very tempting, because so much simpler, to regard a human act as a solid nugget, without internal structure, and to try to trace its moral value to the shape of its surface, as it were. The trolley problem seems to me to bring home that that will not do.)

I said earlier that there seem to me to be two crucial facts about what the bystander does if he proceeds in *Bystander at the Switch*. In the first place, he saves his five by making something that threatens them instead threaten the one. And second, he does not do that by means which themselves constitute infringements of any right of the one's.

Let us begin with the first.

If the surgeon proceeds in *Transplant*, he plainly does not save his five by making something that threatens them instead threaten one. It is organ-failure that threatens his five, and it is not *that* which he makes threaten the young man if he proceeds.

Consider another of Mrs. Foot's cases, which I shall call *Hospital*.

> Suppose [Mrs. Foot says] that there are five patients in a hospital whose lives could be saved by the manufacture of a certain gas, but that this will inevitably release lethal fumes into the room of another patient whom for some reason we are unable to move.[13]

Surely it would not be permissible for us to manufacture the gas.

In *Transplant* and *Hospital*, the five at risk are at risk from their ailments, and this might be thought to make a difference. Let us by-pass it. In a variant on *Hospital*—which I shall call *Hospital'*—all six patients are convalescing. The five at risk are at risk, not from their ailments, but from the ceiling of their room, which is about to fall on them. We can prevent this by pumping on a ceiling-support-mechanism; but doing so will inevitably release lethal fumes into the room of the sixth. Here too it is plain we may not proceed.

Contrast a case in which lethal fumes are being released by the heating

---

12. It may be worth stressing that what I suggest calls for attention is not (as some construals of "double effect" would have it) whether the agent's killing of the one is his means to something, and not (as other construals of "double effect" would have it) whether the death of the one is the agent's means to something, but rather what are the means by which the agent both kills and saves. For a discussion of "the doctrine of double effect," see P. FOOT, *supra* note 1.

13. *Id.* at 29. As Mrs. Foot says, we do not *use* the one if we proceed in *Hospital*. Yet the impermissibility of proceeding in *Hospital* seems to have a common source with the impermissibility of operating in *Transplant*, in which the surgeon *would* be using the one whose parts he takes for the five who need them. This is my second reason for thinking that an appeal to the fact that the surgeon would be using his victim is an over-simple way of taking account of the means he would be employing for the saving of his five. *See supra* note 7.



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

system in the basement of a building next door to the hospital. They are headed towards the room of five. We can deflect them towards the room of one. Would that be permissible? I should think it would be—the case seems to be in all relevant respects like *Bystander at the Switch.*

In *Bystander at the Switch*, something threatens five, and if the agent proceeds, he saves the five by making that very thing threaten the one instead of the five. That is not true of the agents in *Hospital'* or *Hospital* or *Transplant.* In *Hospital'*, for example, what threatens the five is the ceiling, and the agent does not save them by making *it* threaten the one, he saves them by doing what will make something wholly different (some lethal fumes) threaten the one.

Why is this difference morally important? Other things being equal, to kill a man is to infringe his right to life, and we are therefore morally barred from killing. It is not enough to justify killing a person that if we do so, five others will be saved: To say that if we do so, five others will be saved is merely to say that utility will be maximized if we proceed, and that is not by itself sufficient to justify proceeding. Rights trump utilities. So if that is all that can be said in defense of killing a person, then killing that person is not permissible.

But that five others will be saved is not all that can be said in defense of killing in *Bystander at the Switch.* The bystander who proceeds does not merely minimize the number of deaths which get caused: He minimizes the number of deaths which get caused by something that already threatens people, and that will cause deaths whatever the bystander does.

The bystander who proceeds does not make something be a threat to people which would otherwise not be a threat to anyone; he makes be a threat to fewer what is already a threat to more. We might speak here of a "distributive exemption," which permits arranging that something that will do harm anyway shall be better distributed than it otherwise would be—shall (in *Bystander at the Switch*) do harm to fewer rather than more. Not just any distributive intervention is permissible: It is not in general morally open to us to make one die to save five. But other things being equal, it is not morally required of us that we let a burden descend out of the blue onto five when we can make it instead descend onto one.

I do not find it clear why there should be an exemption for, and only for, making a burden which is descending onto five descend, instead, onto one. That there is seems to me very plausible, however. On the one hand, the agent who acts under this exemption makes be a threat to one something that is *already* a threat to more, and thus something that will do harm *whatever* he does; on the other hand, the exemption seems to allow those acts which intuition tells us are clearly permissible, and to rule out those acts which intuition tells us are clearly impermissible.



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

## VII.

More precisely, it is not morally required of us that we let a burden descend out of the blue onto five when we can make it instead descend onto one *if* we can make it descend onto the one by means which do not themselves constitute infringements of rights of the one.

Consider a case—which I shall call *Fat Man*—in which you are standing on a footbridge over the trolley track. You can see a trolley hurtling down the track, out of control. You turn around to see where the trolley is headed, and there are five workmen on the track where it exits from under the footbridge. What to do? Being an expert on trolleys, you know of one certain way to stop an out-of-control trolley: Drop a really heavy weight in its path. But where to find one? It just so happens that standing next to you on the footbridge is a fat man, a really fat man. He is leaning over the railing, watching the trolley; all you have to do is to give him a little shove, and over the railing he will go, onto the track in the path of the trolley. Would it be permissible for you to do this? Everybody to whom I have put this case says it would not be. But why?

Suppose the agent proceeds. He shoves the fat man, thereby toppling him off the footbridge into the path of the trolley, thereby causing him to be hit by the trolley, thereby killing him—but saving the five on the straight track. Then it is true of this agent, as it is true of the agent in *Bystander at the Switch*, that he saves his five by making something which threatens them instead threaten one.

But *this* agent does so by means which themselves constitute an infringement of a right of the one's. For shoving a person is infringing a right of his. So also is toppling a person off a footbridge.

I should stress that doing these things is infringing a person's rights even if doing them does not cause his death—even if doing them causes him no harm at all. As I shall put it, shoving a person, toppling a person off a footbridge, are *themselves* infringements of rights of his. A theory of rights ought to give an account of what makes it be the case that doing either of these things is itself an infringement of a right of his. But I think we may take it to be a datum that it is, the job which confronts the theorist of rights being, not to establish that it is, but rather to explain why it is.

Consider by contrast the agent in *Bystander at the Switch*. He too, if he proceeds, saves five by making something that threatens them instead threaten one. But the means he takes to make that be the case are these: Turn the trolley onto the right-hand track. And turning the trolley onto the right-hand track is not *itself* an infringement of a right of anybody's. The agent would do the one no wrong at all if he turned the trolley onto the right-hand track, and by some miracle the trolley did not hit him.



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

We might of course have imagined it not necessary to shove the fat man. We might have imagined that all you need do to get the trolley to threaten him instead of the five is to wobble the handrail, for the handrail is low, and he is leaning on it, and wobbling it will cause him to fall over and off. Wobbling the handrail would be impermissible, I should think—no less so than shoving. But then there is room for an objection to the idea that the contrast I point to will help explain the moral differences among these cases. For it might be said that if you wobble the handrail, thereby getting the trolley to threaten the one instead of the five, then the means you take to get this to be the case are just these: Wobble the handrail. But doing that is not *itself* an infringement of a right of anybody's. You would do the fat man no wrong at all if you wobbled the handrail and no harm came to him in consequence of your doing so. In this respect, then, your situation seems to be exactly like that of the agent in *Bystander at the Switch.* Just as the means he would be taking to make the trolley threaten one instead of five would not constitute an infringement of a right, so also would the means you would be taking to make the trolley threaten one instead of five not constitute an infringement of a right.

What I had in mind, however, is a rather tighter notion of "means" than shows itself in this objection. By hypothesis, wobbling the handrail will cause the fat man to topple onto the track in the path of the trolley, and thus will cause the trolley to threaten him instead of the five. But the trolley will not threaten him instead of the five unless wobbling the handrail does cause him to topple. Getting the trolley to threaten the fat man instead of the five *requires* getting him into its path. You get the trolley to threaten him instead of them by wobbling the handrail only if, and only because, by wobbling the handrail you topple him into the path of the trolley.

What I had in mind, then, is a notion of "means" which comes out as follows. Suppose you get a trolley to threaten one instead of five by wobbling a handrail. The means you take to get the trolley to threaten the one instead of the five include wobbling the handrail, *and* all those further things that you have to succeed in doing by wobbling the handrail if the trolley is to threaten the one instead of the five.

So the means by which the agent in *Fat Man* gets the trolley to threaten one instead of five include toppling the fat man off the footbridge; and doing that is itself an infringement of a right of the fat man's. By contrast, the means by which the agent in *Bystander at the Switch* gets the trolley to threaten one instead of five include no more than getting the trolley off the straight track onto the right-hand track; and doing that is not itself an infringement of a right of anybody's.



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

## VIII.

It is arguable, however, that what is relevant is not that toppling the fat man off the footbridge is itself an infringement of *a* right of the fat man's but rather that toppling him off the footbridge is itself an infringement of a particularly stringent right of his.

What I have in mind comes out in yet another variant on *Bystander at the Switch*. Here the bystander must cross (without permission) a patch of land that belongs to the one in order to get to the switch; thus in order to get the trolley to threaten the one instead of five, the bystander must infringe a right of the one's. May he proceed?

Or again, in order to get the switch thrown, the bystander must use a sharply pointed tool, and the only available sharply pointed tool is a nailfile that belongs to the one; here too the bystander must infringe a right of the one's in order to get the trolley to threaten the one instead of five. May he proceed?

For my own part, I do not find it obvious that he may. (Remember what the bystander will be doing to the one by throwing that switch.) But others tell me they think it clear the bystander may proceed in such a case. If they are right—and I guess we should agree that they are—then that must surely be because the rights which the bystander would have to infringe here are minor, trivial, non-stringent—property rights of no great importance. By contrast, the right to not be toppled off a footbridge onto a trolley track is on any view a stringent right. We shall therefore have to recognize that what is at work in these cases is a matter of degree: If the agent must infringe a stringent right of the one's in order to get something that threatens five to threaten the one (as in *Fat Man*), then he may not proceed, whereas if the agent need infringe no right of the one's (as in *Bystander at the Switch*), or only a more or less trivial right of the one's (as in these variants on *Bystander at the Switch*), in order to get something that threatens five to threaten the one, then he may proceed.

Where what is at work is a matter of degree, it should be no surprise that there are borderline cases, on which people disagree. I confess to having been greatly surprised, however, at the fact of disagreement on the following variant on *Bystander at the Switch*:

> The five on the straight track are regular track workmen. The right-hand track is a dead end, unused in ten years. The Mayor, representing the City, has set out picnic tables on it, and invited the convalescents at the nearby City Hospital to have their meals there, guaranteeing them that no trolleys will ever, for any reason, be turned onto that track. The one on the right-hand track is a convalescent having his lunch there; it would never have occurred to him



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

> to do so if the Mayor had not issued his invitation and guarantee. The Mayor was out for a walk; he now stands by the switch.[14]

For the Mayor to get the trolley to threaten the one instead of the five, he must turn the trolley onto the right-hand track; but the one has a right against the Mayor that he not turn the trolley onto the right-hand track—a right generated by an official promise, which was then relied on by the one. (Contrast the original *Bystander at the Switch*, in which the one had no such right.) My own feeling is that it is plain the Mayor may not proceed. To my great surprise, I find that some people think he may. I conclude they think the right less stringent than I do.

In any case, that distributive exemption that I spoke of earlier is very conservative. It permits intervention into the world to get an object that already threatens death to those many to instead threaten death to these few, but only by acts that are not themselves gross impingements on the few. That is, the intervenor must not use means that infringe stringent rights of the few in order to get his distributive intention carried out.

It could of course be argued that the fact that the bystander of the original *Bystander at the Switch* makes threaten the one what already threatens the five, and does so by means that do not themselves constitute infringements of any right of the one's (not even a trivial right of the one's), shows that the bystander in that case infringes no right of the one's at all. That is, it could be argued that we have here that independent ground for saying that the bystander does not infringe the one's right to life which I said would be needed by a friend of (ii).[15] But I see nothing to be gained by taking this line, for I see nothing to be gained by supposing it never permissible to infringe a right; and something is lost by taking this line, namely the possibility of viewing the bystander as doing the one a wrong if he proceeds—albeit a wrong it is permissible to do him.

## IX.

What counts as "*an* object which threatens death"? What marks one threat off from another? I have no doubt that ingenious people can construct cases in which we shall be unclear whether to say that if the agent proceeds, he makes threaten the one the very same thing as already threatens the five.

---

14. Notice that in this case too the agent does not *use* the one if he proceeds. (This case, along with a number of other cases I have been discussing, comes from Thomson, *Killing, Letting Die, and the Trolley Problem*, 59 THE MONIST 204 (1976). Mrs. Thomson seems to me to have been blundering around in the dark in that paper, but the student of this problem may possibly find some of the cases she discusses useful.).

15. *See supra* text accompanying notes 9–11.



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

Moreover, which are the interventions in which the agent gets a thing that threatens five to instead threaten one by means that themselves constitute infringements of stringent rights of the one's? I have no doubt that ingenious people can construct cases in which we shall all be unclear whether to say that the agent's means do constitute infringements of stringent rights—and cases also in which we shall be unclear whether to say the agent's means constitute infringements of any rights at all.

But it is surely a mistake to look for precision in the concepts brought to bear to solve this problem: There isn't any to be had. It would be enough if cases in which it seems to us unclear whether to say "same threat," or unclear whether to say "non-right-infringing-means," also seemed to us to be cases in which it is unclear whether the agent may or may not proceed; and if also coming to see a case as one to which these expressions do (or do not) apply involves coming to see the case as one in which the agent may (or may not) proceed.

## X.

If these ideas are correct, then we have a handle on anyway some of the troublesome cases in which people make threats. Suppose a villain says to us "I will cause a ceiling to fall on five unless you send lethal fumes into the room of one." Most of us think it would not be permissible for us to accede to this threat. Why? We may think of the villain as part of the world around the people involved, a part which is going to drop a burden on the five if we do not act. On this way of thinking of him, nothing *yet* threatens the five (certainly no ceiling as yet threatens them) and a fortiori we cannot save the five by making what (already) threatens them instead threaten the one. Alternatively, we may think of the villain as himself a threat to the five. But sending the fumes in is not making *him* be a threat to the one instead of to the five. The hypothesis I proposed, then, yields what it should: We may not accede.

That is because the hypothesis I proposed says nothing at all about the source of the threat to the five. Whether the threat to the five is, or is caused by, a human being or anything else, it is not permissible to do what will kill one to save the five except by making what threatens the five itself threaten the one.

By contrast, it seems to me very plausible to think that if a villain has started a trolley towards five, we may deflect the trolley towards one—other things being equal, of course. If a trolley is headed towards five, and we can deflect it towards one, we *may*, no matter who or what caused it to head towards the five.

I think that these considerations help us in dealing with a question I drew attention to earlier. Suppose a villain says to us "I will cause a



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

ceiling to fall on five unless you send lethal fumes into the room of one." If we refuse, so that he does what he threatens to do, then he surely does something very much worse than we would be doing if we acceded to his threat and sent the fumes in. If we accede, we do something misguided and wrongful, but not nearly as bad as what he does if we refuse.

It should be stressed: The fact that he will do something worse if we do not send the fumes in does not entail that we ought to send them in, or even that it is permissible for us to do so.

How after all could that entail that we may send the fumes in? The fact that we would be saving five lives by sending the fumes in does not itself make it permissible for us to do so. (Rights trump utilities.) How could adding that the taker of those five lives would be doing what is worse than we would tip the balance? If we may not infringe a right of the one in order to save the five lives, it cannot possibly be thought that we may infringe the right of that one in order, not merely to save the five lives, but to make the villain's moral record better than it otherwise would be.

For my own part, I think that considerations of motives apart, and other things being equal, it does no harm to say that

(II) Killing five is worse than killing one

is, after all, true. *Of course* we shall then have to say that assessments of which acts are worse than which do not by themselves settle the question of what is permissible for a person to do. For we shall have to say that, despite the truth of (II), it is not the case that we are required to kill one in order that another person shall not kill five, or even that it is everywhere permissible for us to do this.

What is of interest is that what holds inter-personally also holds intra-personally. I said earlier that we might imagine the surgeon of *Transplant* to have caused the ailments of his five patients. Let us imagine the worst: He gave them chemical X precisely in order to cause their deaths, in order to inherit from them. Now he repents. But the fact that he would be saving five lives by operating on the one does not itself make it permissible for him to operate on the one. (Rights trump utilities.) And if he may not infringe a right of the one in order to save the five lives, it cannot possibly be thought that he may infringe the right of that one in order, not merely to save the five lives, but to make his own moral record better than it otherwise would be.

Another way to put the point is this: Assessments of which acts are worse than which have to be directly relevant to the agent's circumstances if they are to have a bearing on what he may do. If A threatens to kill five



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms

unless B kills one, then although killing five is worse than killing one, these are not the alternatives open to B. The alternatives open to B are: Kill one, thereby forestalling the deaths of five (and making A's moral record better than it otherwise would be), or let it be the case that A kills five. And the supposition that it would be worse for B to choose to kill the one is entirely compatible with the supposition that killing five is worse than killing one. Again, the alternatives open to the surgeon are: Operate on the one, thereby saving five (and making the surgeon's own moral record better than it otherwise would be), or let it be the case that he himself will have killed the five. And the supposition that it would be worse for the surgeon to choose to operate is entirely compatible with the supposition that killing five is worse than killing one.

On the other hand, suppose a second surgeon is faced with a choice between here and now giving chemical X to five, thereby killing them, and operating on, and thereby killing, only one. (It taxes the imagination to invent such a second surgeon, but let that pass. And compare *Trolley Driver*.) Then, other things being equal, it does seem he may choose to operate on the one. Some people would say something stronger, namely that he is required to make this choice. Perhaps they would say that

> (II′) If a person is faced with a choice between doing something ***here and now*** to five, by the doing of which he will kill them, and doing something else ***here and now*** to one, by the doing of which he will kill only the one, then (other things being equal) he ought to choose the second alternative rather than the first

is a quite general moral truth. Whether or not the second surgeon is morally required to make this choice (and thus whether or not (II′) is a general moral truth), it does seem to be the case that he may. But this did seem puzzling. As I put it: Why should the present tense matter so much?

It is plausible to think that the present tense matters because the question for the agent at the time of acting is about the present, viz., "What may I here and now do?," and because that question is the same as the question "Which of the alternatives here and now open to me may I choose?" The alternatives now open to the second surgeon are: kill five or kill one. If killing five is worse than killing one, then perhaps he ought to, but at any rate he may, kill the one.



This content downloaded from
12.13.248.126 on Wed, 14 Sep 2022 20:58:25 UTC
All use subject to https://about.jstor.org/terms