

Lecture I. Who Turned the Trolley?
Lecture II. How Was the Trolley Turned?

F. M. KAMM

THE TANNER LECTURES ON HUMAN VALUES

Delivered at

University of California, Berkeley
March 19–21, 2013

F. M. KAMM is Littauer Professor of Philosophy and Public Policy at the Harvard Kennedy School and a professor of philosophy in the Department of Philosophy, Harvard University. She is the author of *Creation and Abortion; Morality, Mortality*, vols. 1 and 2; *Intricate Ethics*; *Ethics for Enemies: Terror, Torture, and War*; *The Moral Target: Aiming at Right Conduct in War and Other Conflicts*; *Bioethical Prescriptions: To Create, End, Choose, and Improve Lives*; and numerous articles on normative ethical theory and practical ethics. She is a fellow of the American Academy of Arts and Sciences.

LECTURE I.
WHO TURNED THE TROLLEY?

I

Some philosophers, known as act consequentialists, believe that insofar as the greater number of people surviving is the greater good, it is always right for anyone to do what brings about their survival. Nonconsequentialists deny that it is always right to do whatever brings about the greater good, though they need not deny that consequences matter in some ways to the rightness of acts. As support for their view, nonconsequentialists have cited the Transplant Case (in figure 1): If a doctor kills a single healthy person, she will be able to use his organs to produce what we shall assume is the greater good, saving five fatally ill patients, each of whom is equal in all morally relevant respects to the one healthy person. Yet, intuitively, it seems impermissible for the doctor to kill the one to save the five.

However, philosopher Philippa Foot noted that even many nonconsequentialists think it is permissible to kill one to save five in the following case: A driver is driving a trolley (which Foot called a tram) when it becomes clear that it is headed toward killing five people on one track and cannot brake. It can only be stopped from killing the five by the driver redirecting it away from them onto another track where it will kill one different person, who is equal in all morally relevant respects to each of the five.¹ (This is represented in figure 2, with the trolley symbolized by the short horizontal line. All diagrams should be understood to allow the trolley to be moved back from where it is and onto a different track.) Some nonconsequentialists find it mysterious that it is permissible to kill in this case but not in Transplant, even though in both it is a question of killing one to save five, and they think that explaining this mystery just is the Trolley Problem. (Indeed, when Judith Thomson first introduced the title “The Trolley Problem,” she applied it to the question of how to distinguish the permissibility of the driver turning his trolley in Foot’s case from the impermissibility of killing in Transplant.)² However, this suggestion about what the Trolley Problem is should not be understood so narrowly that the problem concerns only these two cases rather than all cases that are structurally similar. For example, one might face the question of whether it is permissible to redirect the trolley off a track and onto a slope leading to someone far away from any track, even though one may not generally kill someone to stop a threat to others. Similarly, the question might arise of whether it is permissible to redirect a

Redirect	Topple to Stop Trolley
Kill vs kill	
Kill vs mere bystander let die	
Kill vs killer let die	

flood or unmanned missile headed to a large city with many people toward a country town with few people in it, even if one may not kill people in some other way to stop the threat to the large city. We might also face an analogous problem when lives are not at stake but only significant harm to either a larger or a smaller group of people. For convenience, I shall use the Transplant and Trolley Cases since they exemplify the general problem.

Even more broadly, we could see the Trolley Problem, understood as I have so far described it, as presenting a challenge to nonconsequentialists who, in some rough way, think there is what is called a side constraint on harming nonthreatening people to produce greater goods. A side constraint is not merely a factor that weighs against producing the greater good; it is supposedly a factor that has priority over producing the greater good. That is, we may produce the greater good only if doing so is consistent with respecting that constraint. One view is that side constraints are individual rights giving rise to correlative duties owed to the right bearer by others and that they reflect fundamental aspects of our conception of persons and their status, which are the background to pursuing persons' good. The challenge is to explain exactly what the side constraint on harming amounts to, and what its form is, if it does not exclude turning the trolley and thus killing one person to save five others.

On the basis of what I have said so far, we can see why the cases that have been the focus of attention in Trolley Problem discussions seem artificial and unrealistic. They are specifically constructed, like scientific experiments, to distinguish among and test theories and principles (for example, consequentialist versus nonconsequentialist theories) because one theory or principle would imply the permissibility of conduct that the other theory or principle would deny. Using our intuitive judgments about which implications for cases are correct helps us decide among, and also revise, theories and principles. (It would not help us with these goals to do as some recommend and spend less time thinking about unrealistic

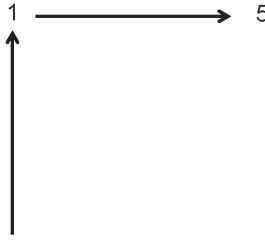


FIGURE 1. Transplant Case: Kill one person for organs to save five.

cases involving hard choices such as letting five die or killing one, and more time making sure that in real life we always have other options available.)

In taking up at least part of the challenge of explaining what the non-consequentialist side constraint on harming amounts to, there are two general questions with which we shall deal in these lectures. First, whether there is a moral difference between killing others and only letting them die such that someone who would otherwise kill many people may permissibly kill fewer other people instead, but someone who would otherwise only let many people die may not kill fewer other people instead. A second question is whether there is a constraint on harming people for the sake of producing a greater good only when we come to harm them in some ways but not in other ways. In other words, is coming to harm them in some ways consistent with proper respect for persons but coming to harm them in others ways not?

These questions are analogous to those typically asked by detectives in mystery cases when someone has been killed, namely, *who* did it and *how* was it done? Like prosecutors, we can also be concerned that the answers to these questions may bear on whether foul play has occurred.

This lecture will focus on the “who” question and lecture 2 on the “how” question. In particular, the first part of this lecture will discuss various elements and proposed solutions of the Trolley Problem as I have so far characterized it. It will then consider an alternative view of what the Trolley Problem is. The second part will deal with some recent views of philosopher Judith Thomson, expressed in her 2008 paper “Turning the Trolley,”³ about how a moral distinction between killing and letting die may eliminate what she now thinks of as the Trolley Problem. The third part will consider a different argument for the conclusion that Thomson favors, but it ultimately attempts to resurrect the Trolley Problem and help defend a position that Thomson herself once held in earlier work.

II

A

Understood as the contrast between Transplant and Trolley, many proposals have been offered to solve the Trolley Problem. One relies on a view known as the Doctrine of Double Effect (DDE), which implies that it is impermissible to kill in Transplant because the doctor *intends* the death of the single person as a means of saving the five. The DDE holds that it is wrong to intend an evil such as a person's death, even as a means to a greater good. In Foot's case, by contrast, the death of the one is only a foreseen side effect of turning the trolley, and according to the DDE merely foreseen harms can be outweighed by producing a greater good.

B

Responding to this proposal, Foot argued that, intuitively, it is often wrong to do what will merely foreseeably cause someone's death in order to maximize lives saved. For example, she said, it seems impermissible for a doctor to use a gas that will save five patients if it is foreseen that the gas will kill one immovable bystander as a mere side effect. Similarly, it seems to me impermissible to set off a bomb that will stop the trolley from hitting the five when a piece of the bomb will kill a bystander as a side effect. (This is the Bomb Trolley Driver Case, as illustrated in figure 3.) Furthermore, it has been argued by some that contrary to the DDE, intending a harm cannot itself make an act impermissible, though it can affect the moral worth of the act.⁴ Here is an example involving a trolley: The driver could permissibly turn the trolley from five even if he did so only because he intended to kill the one person on the other track who is his enemy (Bad Man Trolley Case). Hence, it seems, the DDE fails to explain why killing is permissible in Foot's case but not in Transplant.

III

A

Foot's own proposal is that the driver may redirect the trolley because he is choosing between a negative duty not to kill five and a negative duty not to kill one, and he should kill fewer rather than a greater number. By contrast, the doctor in Transplant is choosing between letting five die and killing one, and the negative duty not to kill takes precedence over the positive duty to aid. Hence, the distinction between negative and positive duties, in these cases represented by not killing and not letting die, is held to

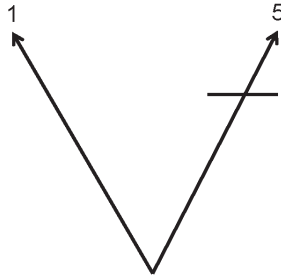


FIGURE 2. Trolley Driver Two-Option Case: Driver can kill five or turn trolley to one.

explain and justify the nonconsequentialist intuitive judgments in Foot's case and in Transplant.

One possible objection to Foot's proposal is that it could be unclear whether the trolley driver is in the process of doing what will kill the five or rather is only in a position to let what will kill them—namely, the out-of-control trolley—occur. If the latter were true, he may become the killer of the five in virtue of having first started the trolley and choosing to let it continue, but at the time of deciding whether to turn it perhaps he would only be letting the five die if he did not redirect. Judith Thomson agrees with Foot that the driver would be killing the five if he does not redirect. In her 2008 article, she imagines a case where a car driver “suddenly sees five people on the street ahead of him, but his brakes fail: he cannot stop his car, he can only continue onto the street ahead or steer to the right (killing one) or steer to the left (killing himself). If he doesn't steer to one or the other side, if he simply takes his hands off the wheel, he runs the five down and kills them. He cannot at all plausibly insist that he merely lets them die. So similarly for the trolley driver.”⁵

Hence, both Foot and Thomson see Foot's case as like one in which the driver is at a cross point and must decide whether to turn the trolley toward five or toward one. Let us assume that this is true in Foot's case, though we shall return to this issue in discussing other cases below.

B

Here is another objection to Foot's proposal that someone who would otherwise kill, but not merely let die, may kill fewer other people. The objection is that her proposal would permit too much. To see this, suppose the driver could stop the trolley from hitting the five only by pressing a button that causes a device to topple a fat man standing on a bridge so that

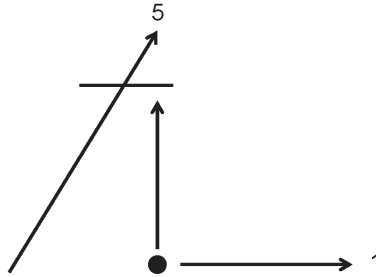


FIGURE 3. Bomb Trolley Driver Case: Driver can kill five or set off bomb to stop trolley; bomb kills one.

he falls in front of the trolley. His weight would stop the trolley, though he would be killed by it. (Call this the Driver Topple Case, shown in figure 4.)⁶ Intuitively, toppling the fat man seems as impermissible as killing in Transplant, even if toppling, like redirecting, would involve the driver killing one rather than killing five by mechanical means. (My case involves the use of a mechanical device because some have proposed that it is only an “up close and personal” push of the fat man that is impermissible, and I disagree with this.) Thomson introduced a case in which a *bystander*, not the driver, can stop the trolley only by pushing the fat man, and she concluded that his doing so was impermissible.⁷ However, she did not consider whether the *driver* may push the fat man or topple him using a mechanical device when this is his only way to stop the trolley. In addition, as already suggested by figure 3, I think the driver may not set a bomb to stop the trolley when a piece of the bomb would kill another person as a side effect.

Those who will otherwise impermissibly kill the five may not do just anything to minimize the number that they kill, even when doing so would be the strict alternative at a specific time to killing the five. Suppose the driver’s taking his hands off the steering mechanism would be his killing the five (as Thomson says) and his keeping his hands on it would involve his pressing the button that topples the fat man whose fall stops the trolley. Here, toppling the one to stop the trolley would be the alternative to one way of killing five, but it seems impermissible. Foot’s failure to consider the limits on the driver is one problem with her analysis, I think.

Given the limits on the ways in which the driver may bring about others’ deaths, some might think that the Trolley Problem is to explain why it is permissible to kill some people in some ways but not in other ways as the

alternative to killing a greater number of different people, even if it is not permissible to kill someone as the alternative to merely letting others die.

C

Another challenge to the adequacy of Foot's killing-versus-letting-die explanation of the difference between Transplant and her trolley case is that it would permit too little, ruling out some permissible acts. This was the point of Thomson's brilliant Bystander Two-Option Case in which the driver is unable to do anything to stop the trolley, but a bystander (Mr. X) can press a switch and redirect the trolley.⁸ (This case is the same as represented in figure 2 except that a bystander, not a driver, would turn the trolley.) The bystander's doing so would involve his killing one rather than *letting five die*. If it is permissible for him to redirect the trolley, the distinction between otherwise killing and otherwise letting die will not explain why it is permissible for the driver to kill one in Foot's trolley case but impermissible for the doctor to kill one in Transplant when, like the bystander, the doctor too would otherwise let five die. Possibly, the killing-letting die distinction could still explain why the driver who would otherwise kill the five might have a duty to kill fewer people in a permissible way, while a bystander who would merely let five die has the option but not a duty to permissibly kill one other person.

Now consider a case in which the driver himself would be choosing between killing one and letting five die. Suppose the driver is thrown from the trolley and so is not on the threatening entity but is near a switch at the side of the road with which he could redirect the trolley. (Call this the Bystanding Driver Two-Option Case.) I believe it is permissible for the driver to redirect the trolley in this case. If so, the fact that at a given time someone would be choosing to kill one person rather than to let five others die will not rule out killing.

In many possible cases, it seems clear that sometimes an agent who starts a threat to some people and then separately does not save them from the threat because, for example, the only way to save them would require killing another person will have let the people die at the time he does not save them. This is so even though in virtue of this and his past act he will have killed the person.⁹ Indeed, Thomson once discussed a version of Transplant like this.¹⁰ She imagined that the five in Transplant suffered from organ failures caused by the very doctor who contemplates taking organs from his healthy patient to save his own victims. She concluded

that at the time when he would act, this culpable doctor faced a choice between killing one and letting the five die and that he had no more right than any other doctor to kill in Transplant, even though he would become the killer of the five if he did not save them.

Suppose that unlike this culpable doctor, the driver in the Bystanding Driver Case is permitted to avoid having killed five by redirecting to one other person, even though he is not permitted to topple the fat man to stop the trolley any more than the culpable doctor may kill the one person to save his five victims. Then it might still be suggested that bystanders who will (as I shall say) *merely* let five die and will not also thus become their killers may not kill by redirecting the trolley. But, of course, it was the point of Thomson's original mere Bystander Case to argue against this last claim by suggesting that a mere bystander may also redirect the trolley. However, she also claimed that the mere bystander may not topple the fat man any more than the doctor in the original Transplant Case may kill one patient to save five.

Figure 5 displays some of the distinctions we have been discussing.

In her 2008 article "Turning the Trolley," Thomson claims that the question of why a bystander may permissibly redirect the trolley rather than let five die but not topple the fat man to stop the trolley is (the whole of) the Trolley Problem.¹¹ She refers to the question of why a driver who is on the trolley driving it and would kill the five may redirect the trolley but the doctor may not kill in the original version of Transplant as "Mrs. Foot's problem." It may seem surprising that she says Mrs. Foot's problem is *not* the (or part of the) Trolley Problem given that when Thomson introduced the title, she applied it to Mrs. Foot's problem (as quoted in note 2). However, in a later article, she did switch her use of the title "Trolley Problem" so that it applied only to the contrast between the bystander turning the trolley and his toppling of the fat man.¹²

It seems to me that the Trolley Problem should not be conceived as only the question of why a mere bystander may redirect the trolley but not topple the fat man to stop it. This is, in part, because the same problem of redirecting being permissible but toppling not being permissible arises, I claim, for the trolley driver. This is so whether he would otherwise kill the five at the time or would let them die, as in the Bystanding Driver Case. Both Thomson and Foot fail to consider what the driver may not do and what the bystanding driver may do. In addition, I believe that Foot's original problem is (as it was originally considered by Thomson to be) at

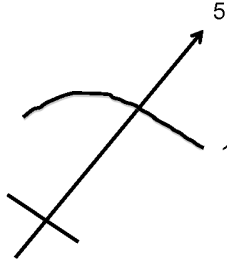


FIGURE 4. Driver Trolley Case: Driver can kill five or topple man in front of trolley, killing him to save five.

least part of the Trolley Problem. It may just be best to call it and all the “cousin cases” of that original problem parts of the Trolley Problem.

D

The most striking claim offered by Thomson in her 2008 article is that she now believes that it is not permissible for a mere bystander to redirect the trolley. Indeed, in that article Thomson reversed herself, claiming that it is not any more permissible for the bystander to redirect the trolley toward the one than to topple the fat man to stop the trolley,¹³ and, therefore she, a founding mother of the Trolley Problem, believes there is no Trolley Problem (given her second view of what the problem is). On these grounds, she further claims that her original Bystander Case cannot be used as an objection to Foot’s proposal for distinguishing Foot’s Trolley Case from Transplant, namely, that killing one is the permissible alternative to killing five but not to letting five die (and so the driver may redirect, but the doctor may not kill). Indeed, she endorses Foot’s proposal for distinguishing Foot’s Trolley Case and Transplant.

In the next part of the lecture, I wish to consider whether Thomson’s reversal on the permissibility of a mere bystander redirecting the trolley is right (or whether we should defend something like Thomson’s original position against her revised one).¹⁴ However, it is first important to re-emphasize that if being in the position of letting five die rather than killing one rules out turning the trolley, then the bystanding driver will also not be permitted to turn the trolley. If, as I believe, he is permitted to do this, Foot’s proposal might be minimally revised to allow it on the grounds that “one may not kill rather than merely let others die, but one may kill rather than become the killer of others.” But if a Bystander Driver or even a regular driver is permitted to kill by redirecting but still not permitted

to topple the fat man to stop the trolley killing the five, then at least one component of the broader Trolley Problem will remain, namely, the moral differences in how we come to kill someone.

IV

Let us now put aside the Bystanding Driver Case and consider Thomson's 2008 discussion of her mere Bystander Case.

A

One of her arguments against the permissibility of a mere bystander redirecting is based on what she calls the Bystander's Three-Option Case, in figure 5. By contrast to the ordinary Bystander's Two-Option Case in which the bystander has a choice between letting five die or redirecting the trolley to a track where one different person will be killed by it, in the Bystander's Three-Option Case: "The switch available to this bystander can be thrown in two ways. If he throws it to the right, then the trolley will turn onto the spur of track to the right, thereby killing one workman. If he throws it to the left, then the trolley will turn onto the spur of track to the left. The bystander himself stands on that left-hand spur of track, and will himself be killed if the trolley turns onto it. Or, of course, he can do nothing, letting five workmen die."¹⁵ Concerning her Bystander's Three-Option Case, Thomson first argues that the bystander may not redirect toward the right when he could redirect toward himself as "if A wants to do a certain good deed and can pay what doing it would cost—other things equal—A may do that good deed only if A pays the cost himself." (Call this the Self-Cost Claim.) Second, she argues that since the bystander is not morally required to sacrifice his life to save the five, doing so would be altruism (and supererogatory, I would add). Hence, given that he has permission only to harm himself to save the five when he can harm himself, and he does not want to and need not do so, he should let the five die. When in the Bystander's Two-Option Case there is actually no option of turning the trolley on himself, Thomson says there is still "no way in which he can decently regard himself as entitled to make someone else pay" a cost that he would not pay if he could.¹⁶

Let us consider this first argument before presenting another one that she offers. In saying, in the second step of the argument, that the bystander need not sacrifice his life to save the five, Thomson is relying on another aspect of a nonconsequentialist ethical theory: that one has, in general, no duty to do what will bring about the greatest good at extreme cost to

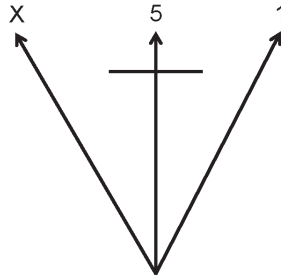


FIGURE 5. Bystander's Three-Option Case: Bystander (X) can let five die or turn to one or turn trolley to self.

oneself, even if one would violate no side constraints in doing so. I agree with this part of her argument. It helps us see that one could raise some of the same issues that Thomson raises by her Three-Option Case by using a two-option case. Suppose, for example, that it would break a bystander's back to turn the trolley away from the five. From a nonconsequentialist point of view, this could be a justification for his not turning it even though the five will die. Suppose that to avoid breaking his back, he permissibly would not turn it. How then, we might ask, can he decently turn the trolley in cases where it would not break his back but will kill one other person?

My first concern with Thomson's argument is that she moves from a three-option case to a conclusion about a two-option case. It might be said that one cannot, in general, move from a conclusion in a three-option case to one in a two-option case (and vice versa). For example, if we have three choices, between letting a murder occur or stopping it by killing the murderer or stopping it by shooting him in the leg, it is impermissible to kill him. However, this does not mean it is impermissible to kill him if it is one's only alternative to letting the murder occur.

However, in these cases, the third option—shooting the murderer in the leg—which is not present in the two-option case, would be taken in the three-option case. In Thomson's argument, the third option of the bystander sacrificing himself—which is missing in the two-option case—would permissibly *not* be taken. So it might be thought that if an option would permissibly not be taken even if it were present, its absence in the two-option case could not affect what may be done, and so if it was impermissible to redirect to the bystander in the three-option case, it would be so in the two-option case.

But this is not so. Whether an option is permissible may depend on what the alternative is even if that alternative would not be taken. For

example, there might be an interaction effect caused by the mere presence of the third option even if it is not taken, and this interaction effect might make the second option impermissible in the three-option case without its being impermissible when the third option is not present. In particular, making a choice between oneself and another might necessitate treating the other as one treats oneself even if one need not treat him as one would treat oneself when there is no such choice to be made.¹⁷ So someone might agree with Thomson that in the three-option case the bystander may not choose to sacrifice the other when he could sacrifice himself, or, more weakly, that he must at least give himself and another equal chances of being sacrificed, yet hold that he may turn the trolley on the one person in the ordinary two-option case.

In fact, I think that in the three-option case there is no special interaction effect between the second and third options, and the bystander need not refrain even in that case from imposing costs on another person that he could, but will not, impose on himself.

My second major concern with Thomson's argument is that in considering her Bystander's Three-Option Case, Thomson first concludes that it is impermissible for the bystander to turn the trolley to the one person when he could turn it to himself instead. Only afterward does she argue that it is permissible for the bystander not to turn it to himself to save the five.¹⁸ Let us consider the first step on its own. Suppose, for argument's sake, that the bystander's turning the trolley toward the one person rather than toward himself would be indecent. This alone may not show that it is impermissible to so turn. Consider a case in which only a very small cost would be necessary to do the good deed of saving the five (for example, throwing them one's spare life preserver), but the bystander is not willing to pay it. From what Thomson says, it seems to follow that it would be impermissible for him to make another person pay the very small cost in order to save the five (for example, throw the other person's spare life preserver when that person also does not want to throw it). But even if it was indecent, it does not seem to be impermissible because refusing to pay the small cost is, presumably, wrong of each person given that five will otherwise die. (Thomson's objection does not seem to be to imposing a cost on someone against his will if he would clearly have a moral duty to pay the cost himself.) Hence, the Self-Cost Claim does not seem to be true.

While I agree that the bystander's not giving up his life to save the five is not in itself impermissible, my point here is that at the very least, one needs to first show that it is permissible for the bystander or the other

person not to make a sacrifice to save the five before one argues for the impermissibility of the bystander imposing the cost on another if he will not pay it himself. If the cost is required of each, a person's failing to make the sacrifice may not show that it is impermissible for him to impose it on another.

Indeed, a very strong general claim might be suggested (Strong Claim 1): If it would be permissible to impose a cost on someone for some end were there no alternative, it need not become impermissible to do so merely because one does not impose the cost on someone else when one should. (This claim is meant to apply whether the person on whom one should impose the cost is oneself or another person.) If this is so, then one cannot show that it is impermissible to impose a cost on another when it is either the only way or one of the ways to achieve a goal simply by showing that one would not impose it on oneself, even if one should impose it on oneself rather than on another. For example, suppose it is permissible to impose on rich people the small cost of saving many lives. However, one morally should impose the cost on the richest of the rich before one imposes it on the less rich. Even if it is morally wrong not to follow this order in deciding on whom to impose, this alone does not make it impermissible to impose the cost on the less rich person. It may, of course, be impermissible to impose certain costs on others. My point is that it is not a sufficient argument for showing this that one will not or would not oneself pay the costs. (Below I will give another reason to believe this.) And if one had an independent argument for the permissibility of turning the trolley on someone, one should not assume that it is defeated by the fact that one would not turn the trolley on oneself even if failing to do so was wrong (which I do not think it is).

As background to my third major concern about Thomson's first argument, notice that some philosophers have been concerned with whether one has the standing to condemn someone's doing the same thing that one has done or is doing, or require that someone else *do* something when one is oneself refusing or has refused to do it, other things being equal. This type of concern might imply that it is indecent or impermissible for the bystander to either condemn someone else for refusing to *volunteer his life* or *require him to volunteer his life* when the bystander is refusing or has refused or would refuse to volunteer his own life in the three-option case, other things being equal. By contrast, Thomson wishes to condemn as indecent and impermissible the bystander's *taking* someone's life without that person's consent when the bystander either will not or would not

volunteer his own. But taking someone's life without his consent involves imposing a cost on another, and this is not the same as demanding that the other person impose the cost on himself altruistically when the bystander will not or would not impose the cost on himself altruistically. It may, of course, be impermissible to impose costs on others without their consent, but I do not think, contrary to Thomson, that a good argument for showing this is that one will not or would not oneself *volunteer* to pay the costs.

I think that sometimes it may be morally permissible to impose costs on others when no one would, or would be required to, impose them on himself (though it is permissible for him to impose them on himself). Here are some examples. We may sometimes draft someone to be a soldier even if we (and also he) would not and need not volunteer for service (though perhaps only a government may do the drafting). I may permissibly present evidence leading to someone's conviction even if I would not and need not bear witness against myself (nor would he, or need he, bear witness against himself). I may permissibly compete with someone causing him to lose his business so that I get funds for a cause, though I would not, and need not, close my own business to get funds for a cause. Getting closer to the Trolley Problem, I may put a shield around the five people, protecting them from the trolley, knowing that when the trolley hits the shield, it will be deflected to one other person, even if I would not and need not put a shield around the five when I know the trolley will be deflected to me. (Notice that in the draft case, resistance by the person imposed on is not [ordinarily] permissible. In the other cases, I think the fact that I may permissibly impose costs on another does not imply that the other may not resist the costs being imposed [at least in certain ways]. So the person threatened by the trolley after it is deflected off the shield may permissibly send it back if he can.)

Hence, it is possible that a bystander's *not being willing to* redirect the trolley toward himself implies only that it is "indecent" of him to require someone else to redirect the trolley toward *himself*; it need not imply that it is indecent for him to redirect the trolley to someone else. Similarly, if it is *permissible* for the bystander not to give up his own life, he should recognize that the other person on the track also has the permission not to give up his own life. Further, if it is permissible for the bystander to impose the loss on the other person, he should recognize that it is also permissible for someone else to impose the loss on the bystander in order to save the five. This is the way in which impartiality is required of us. Similarly, if we

all established a military draft, each is required to expose himself to the risk of being drafted, but this is consistent with no one having to volunteer for service.¹⁹

Certainly, in cases of letting die, rather than killing or otherwise doing what harms someone, this difference between volunteering and imposing (in a broad sense) holds. Suppose that the only way for me to save five others from drowning is to let myself drown (by giving them a life preserver I need and could otherwise use). I could permissibly refuse to let myself drown and let the five drown instead. Yet it would also be permissible for me to let one *other* person drown, though he does not want me to, in order to save five others instead (by giving the five instead of the one a life preserver when I do not need it). I would permissibly impose a cost on that one by not aiding him that I would permissibly not impose on myself. This, of course, is a case of letting someone die, and Thomson believes (as do I) that there is a moral difference between killing someone and letting him die, but to show that a bystander may not redirect the trolley to the other person, more has to be said than that, other things being equal, he should not make someone else pay a cost that he would not and need not pay himself. This factor would also imply that he may not let the one die to instead save the five if he would not and need not let himself die in order to save the five. If one has an independent argument for the permissibility of letting someone die to save others, one cannot show that it is impermissible to do so merely because one would permissibly not let oneself die to save others.

B

Thomson's second argument for the impermissibility of the bystander turning the trolley explicitly does *not* depend on considering what the bystander would or would not do to himself. She notes that even if the bystander would be an altruist if he could, the one person he would kill if he redirected the trolley in a two-option case might not, and need not, consent to be an altruist.²⁰ This, she thinks, makes the mere bystander doing what costs the other person his life impermissible. Whereas the first argument claims that one may not do to another without his consent what one need not and/or would not do to oneself, the second argument claims that one may not do to another what he is not required to do to himself and would not consent to have done to him. If this argument was sufficient in itself to make turning impermissible, the fact that the first argument is problematic would not matter.

So let us consider if it is sufficient. First, as was noted earlier, imposing a cost on someone is not the same as requiring him to act altruistically by either engaging in self-sacrifice or consenting to someone's imposing a cost on him. Furthermore, it is not generally impermissible to do something to save some people when it would lead to one other person suffering a loss he need not and permissibly does not consent to suffer. This is true in the case described earlier where we put a shield around the five, knowing that it will deflect the trolley to the one. The cases mentioned earlier involving harmful competition and bearing witness against another also involve imposing costs on the other person he is not morally required to impose on himself and does not consent to suffer. Further, a person could permissibly not consent to give up his own life or a bystander's lifesaving assistance so that the bystander may save five other people instead, yet it would be permissible for the bystander to help the five and (in a broad sense) impose the loss on the one. If one has an independent argument for the permissibility of leaving someone to die, this argument is not defeated merely by showing that the person need not let himself die. (It is probably crucial to Thomson whether one would be killing rather than not aiding the one person, but I think that her argument is too broad to capture this distinction among cases.)

Hence, for the various reasons I have given, I am not convinced of the truth of either Thomson's *ceteris paribus* claim (which is her Third Principle) that bystander "A must not kill B to save five if he can instead kill himself to save the five," nor of her claim that A must not kill B to save five others if B would permissibly not altruistically sacrifice himself. However, it is important to emphasize that the permissibility of turning the trolley need not imply that the one person who would be hurt by it may not resist this even by turning back the trolley. This is consistent with his having no duty to be an altruist and also implies that it is permissible to resist (what one recognizes is) another agent's permissible act.

C

Now let us consider Thomson's argument for the view that the trolley driver is permitted and even required to redirect to the one person, even though that one does not and need not volunteer to be an altruist or consent to the imposition of the loss. (A sign that he need not consent, I believe, is that he may resist being killed even by sending the trolley back or toward the driver who sends it to him. Hence, he may permissibly resist the driver's act even if it is the driver's duty, and nothing about any duty the driver may have gives the one person a duty to consent to the

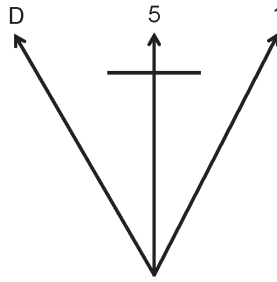


FIGURE 6. Trolley Driver Three-Option Case: Driver can kill five or turn trolley to one or turn trolley so it kills himself.

imposition.) Thomson also thinks that the driver should redirect, even though he would not redirect the trolley in a way that will only kill himself to save the five when he has this third option. (This case is represented in figure 6. Thomson does think that the driver, unlike the bystander, has a duty to kill himself and would be unjust in not killing himself to save the five.²¹ This raises the possibility that she might think it permissible for the mere bystander to redirect the trolley when this will kill only the driver on the trolley.) The driver who will not kill himself or who has no option to do so is permitted to turn and even must turn toward the unwilling one, she says, because the driver would otherwise unjustly kill five who are also unwilling to be killed. By contrast, Thomson says, the mere bystander will have done no injustice if he leaves the five to die, and this makes a significant difference to the permissibility of killing the one.

There are several things that concern me about this additional component of the argument relating to the driver. First, it is possible that a driver who will become the killer of five people because a trolley that he started will, due to factors beyond his control, kill five people need not be required to deliberately do what will kill someone else even if this is by redirecting the trolley. Suppose the reason the trolley driver gave for letting the trolley go on to the five was that he could not deliberately do what will kill someone in order to stop the killing of five others when that will be due in large part to factors beyond his control. I doubt that we would say that this in itself involved his treating the five unjustly, just as we would not say this if he refused to topple the fat man to stop the trolley.

However, let us assume that the trolley driver would be unjust to the five both in killing them and if he does not turn the trolley in the ordinary two-option case. Thomson's first argument claimed that the mere bystander is permitted to not save the five when the cost to him of doing so

would be very high (for example, his life). But what if the cost to the mere bystander personally is very small and furthermore he had promised to save the five? Then his not saving them may also involve some injustice to them. Presumably, in this case Thomson's view would be that it is impermissible for the bystander to turn the trolley, since killing one person on the track is a more serious injustice than breaking a promise to save the five. My point is only that the supposed impermissibility of a mere bystander turning the trolley cannot depend on his committing no injustice if he does not turn.

Of course, in arguing that the driver is permitted to redirect when the bystander is not, Thomson does not merely rely on the fact that the bystander's letting die involves no injustice, whereas killing the five people is an injustice. She also relies on what she calls the Killing Five versus Killing One Principle, which specifically says that A must not kill five if he can kill one instead.²² But this principle does not seem to be true. As I noted earlier, even the driver with his hand on the wheel headed to the five may not set a bomb (with his other hand) that will stop the trolley but also directly kill one other person, and he may not topple the fat man from the bridge in order to stop the trolley from killing the five. The fact that the driver will unjustly kill five people if he does not do what kills one other person, together with the Killing Five versus Killing One Principle, is not enough to free the driver to redirect to the other person if it does not free him to topple the fat man off the bridge or set the bomb as he is driving the trolley toward the five. Thomson says, "But his not steering to the right would itself be unjust; for his only alternative to steering to the right is killing five."²³ But if we could deduce the injustice of the driver not turning the trolley from the fact that his only alternative is to kill five, we could also deduce the injustice of his not toppling the fat man from the fact that his only alternative is to kill five, but it is not unjust of the driver not to topple the fat man to stop the trolley.

In sum, as I have noted several times, Thomson in her 2008 article does not attend to what even the driver is not permitted to do. She also does not consider what the bystander driver may do. This may be why she supports the Killing Five versus Killing One Principle. As I have already said, it may also be why she says that once we accept that the mere bystander may not redirect, the distinction between killing and letting die explains why the driver who would otherwise kill five may redirect the trolley, but the doctor may not kill in Transplant and the bystander who would merely let five die may not topple the fat man. But this explanation would not rule out the driver toppling the fat man to stop the trolley when it is the

only way to stop his driving into the five. This explanation would also rule out the bystander killing rather than letting die. Hence, it seems to me the explanation is not correct. Thomson is wedded to the idea that the driver must be permitted and has a duty to turn the trolley from five to one. Indeed, she is willing to conclude that her reasoning about the mere bystander is wrong if it leads to the view that driver may not turn the trolley.²⁴ In the end, what we may find most surprising about Thomson's discussion is not that she argues that the mere bystander may not redirect the trolley but that her arguments seem to imply that a driver who would otherwise kill five may, at the same time as an alternative, kill in any way at all fewer people in order to stop from killing the five.

It might be suggested that something else, either on its own or in addition to his strong duty not to commit an injustice and kill the five, explains the limited ways in which the driver may kill the one person to save the five (for example, turning but not toppling). This additional factor could be that the driver's duty is to drive a trolley in the best possible way, so when not killing the five requires driving a trolley that kills one other person, he is permitted to kill. By contrast, toppling people or setting off bombs is not a duty of his, and in these respects he is not otherwise freer than a mere bystander. However, if the driver had to topple a rock or set off a bomb that would not hurt anyone, he would be obligated to do these things to save his five potential victims, even though he has no special responsibility for rocks or bombs. It is the permissibility of what he could do that determines his duty to do it rather than to kill the five, not his special duty for driving his trolley that determines permissibility.

If the driver may do these other things besides drive the trolley to stop it from killing the five, then we cannot explain his not having permission to set off the bomb that will kill someone or topple the fat man to stop the trolley by his special responsibility for driving his trolley well. This again suggests that there is some difference in how an agent kills an innocent, nonthreatening person that makes some killings to prevent five being killed permissible and other killings impermissible. A further hypothesis is that this difference marks a distinction between permissible and impermissible killings in general and so for a mere bystander, not just for the driver. If this is so, that the killing would be done by someone who would otherwise kill a greater number of people rather than merely let them die would not be crucial for permissibility.

Indeed, one might be tempted by a strong general claim (Strong Claim 2): that one will otherwise unjustly kill innocents does not itself

give one any *special* permission that others lack to kill other innocent, non-threatening people to prevent the first injustice. (It might, however, give one excuses for killing impermissibly and obligations to kill, in ways permitted to everyone, to prevent oneself from killing others. These excuses and obligations would not be had by mere bystanders.) This strong claim is consistent with the person who would otherwise unjustly kill innocents having other permissions that a mere bystander would lack. For example, he might have permission to sacrifice his wealth to save his potential victims even if this would seriously deprive his family. A bystander might not have such permission to do what seriously deprives his family to save the potential victims.

Strong Claim 2 would also have to be consistent with the fact that having to not kill people may give one a stronger reason for action than merely not letting them be killed by others and that ordinarily having a stronger reason can make acts be permissible that are not permissible for weaker reasons. For example, it might be permissible to take your car without your permission if I have to get someone to the hospital but not if I only want to go for a drive. Suppose, however, that I do not have a duty to take someone to the hospital who needs to go but choose to do so anyway. I think it as permissible for me to take the car as it is for a person who has this duty. It is that someone needs to be saved, not whether I have a duty to save him, that provides a reason that overrides the duty not to interfere with someone else's car. Someone's having a duty to save a person may be part of the stronger personal reason that that person has to act, but it is not necessarily part of the stronger reason that overrides another's right to his car. (There are two different types of stronger reason.) This supports the permissibility of the mere bystander turning the trolley if the driver may turn it, even if the bystander does not have an obligation to do so. That is, the driver has a stronger personal reason for taking advantage of the factors that may make turning permissible, but the factors that make it permissible are present for anyone.

However, here is a case that may be a counterexample to Strong Claim 2. Suppose a driver finds that his trolley will kill many people if it remains at a crossroad where he has driven it, and he can redirect to the left, where it will kill Jim, or to the right, where it will kill Joe. Then he might toss a coin to decide which single person to kill. But it seems that a mere bystander should not toss a coin to decide whether to redirect a trolley to kill Jim on the left if the alternative is to let it go where it is already headed, toward killing Joe on the right. Is this different from what someone who would

otherwise kill may do? Consider whether a driver whose trolley is already headed toward killing Joe on the right may toss a coin to decide whether to redirect to Jim on the left any more than a mere bystander. He may not be permitted to do this either, I think. Still, this is consistent with anyone who faces a choice between killing Jim or killing Joe sometimes (for example, in the crossroad case) being permitted to kill Jim, while someone who faces a choice between killing Jim or letting Joe die is not permitted to kill Jim. In the light of such possible equal-number-of-people cases, Strong Claim 2 might have to be revised to the following Strong Claim 3: that one will otherwise unjustly kill a greater number of innocents does not itself give one any special permission that others lack to kill fewer innocent, non-threatening people to prevent the first injustice.

There might be indirect evidence that it is the general permissibility of killing only in certain ways that is at work in Trolley Cases, rather than that stronger reasons for not killing than for not letting die make killing permissible only for the driver. But the most direct route to showing that what is at issue in Trolley Cases is the general permissibility of killing only in certain ways is to give a good account of the moral differences between different ways of killing that would help explain why some ways are permitted to mere bystanders as well as to potential killers. This is the question of how the trolley was turned rather than by whom it was turned. It is the question with which Thomson was at one time concerned when she said, "We ought to be looking within killings and savings for the ways in which the agents would be carrying them out."²⁵ I shall not discuss that question in this lecture but will discuss it in lecture 2.

v

I have considered Thomson's arguments for the conclusion that one may turn the trolley only if this is the alternative to one's killing people, not to one's letting them die. However, I believe the strongest argument for this conclusion may be one she does not give. Consider the following Saving-by-Letting-Die Case in the top part of figure 7: Suppose the trolley is on its way to kill eight people. A mere bystander can either turn it away to the left, thereby killing one person on a side track, or turn it to the right, thereby blocking the pathway that he must and will shortly use to save five other people, for whose problems he is not responsible but whom no one else can save. In this case, he faces the choice of saving the eight by either killing one person or doing what (arguably) results in his letting five people die. I believe the bystander should do what blocks the pathway, and

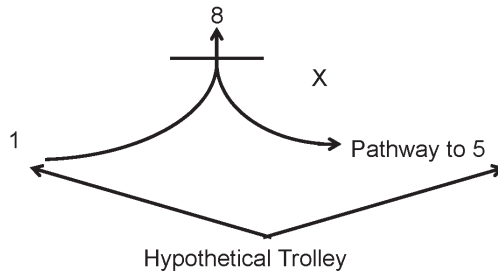


FIGURE 7. Bystander Saving-by-Letting-Die Case.

so do what (arguably) results in his letting five people die rather than in killing one. In this case, it seems, the bystander should not kill someone as the alternative to letting five die. How then could it be permissible for the bystander to kill one person rather than let the five die in Thomson's original Bystander Case?

Indeed, what seems especially puzzling is that while the bystander should turn the trolley from the eight in a way that results in his letting the five die rather than in killing the one, if a trolley had been headed to killing the same five people when they were not soon to die anyway, then, intuitively, it would have been permissible for the bystander to turn it away from them so as to kill the one person. (See the bottom of figure 7.) How can these judgments be consistent? How can it be both permissible to kill one rather than to let five die in the latter case and impermissible to kill one rather than to let five die in the Saving-by-Letting-Die Case?

There are several differences between the ordinary trolley case and the Bystander-Saving-by-Letting-Die Case that might be thought to explain why the judgments are consistent. One is that when the eight are saved by doing what results in letting the five die, the redirected trolley interferes with some means needed (the pathway) to save the five; it does not threaten the five as it threatens the eight or as it would threaten the one. This would contrast with the regular trolley cases in which the trolley threatens the five, not just some means needed to save them. So we should consider another hypothetical two-option trolley case in which a trolley is headed to blocking the pathway that a bystander alone could and would use to save five people; it is not headed to the five themselves. In this case, may the bystander redirect the trolley when it will kill one other person instead? I do not think he may do this.²⁶ Most important, the driver of the trolley headed to block a pathway that he alone would use to save five (who were not threatened by him) should not turn the trolley toward killing one

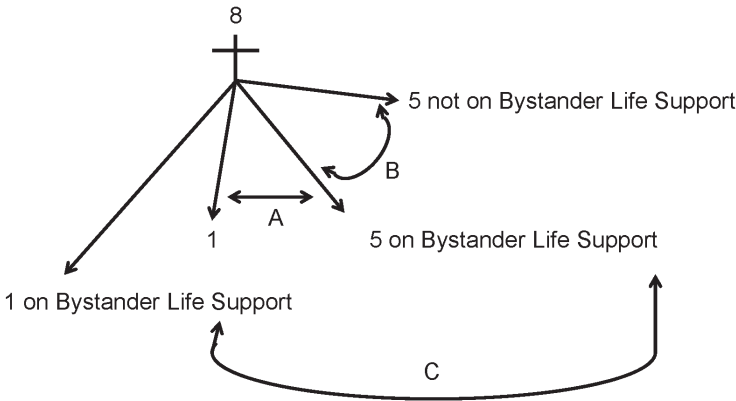


FIGURE 8. Bystander Saving-by-Letting-Die Case (2).

person. Letting some die by allowing or even causing interference with, or destruction of, what is needed to save them (especially when one alone could provide or make use of these means and one is not interfering with others using them), at least intuitively, seems to have different moral significance from letting some die by not stopping what actually will cause their death (as when a trolley is coming at them).

However, I do not think that pointing to a possible distinction between rescuing means and rescuing a greater number of people is a complete answer to issues raised by the Saving-by-Letting-Die Case. Consider the following case: Suppose the bystander has a choice of turning the trolley from the eight either to the left, where it will kill one other person, or to the right, where it will disconnect his privately owned life support machine that is already saving five people's lives. (See figure 8, choice A.) I believe that terminating one's ongoing life support of others in this way is a case of letting the five die, not of killing them. It seems, at least when it is optional for the bystander to be aiding the five, that he should turn the trolley to where it terminates his assistance to the five rather than turn to where it will kill the one person.

If he should do this, it would be, at least in part, because the bystander has the permissible option to actively terminate life support he provides and allow the five to die. (I am ignoring the view that having started aiding some, one is committed to them in a way that one is not to those one is not already aiding, and that is a definitive consideration.) But in the regular Bystander Case, he also has the permissible option to let the five die, so why is it not true then also that he should let the five die rather than

kill the one? Could we account for a difference in these cases as follows: The trolley headed to five in the ordinary Bystander Case is threatening the lives of those who are not already receiving optional life support; they would lose life that they have a claim not to be interfered with by the trolley driver. So they would have a complaint if they were killed by the trolley driver. A trolley redirected by the bystander to the five on the bystander's life support will cause the five to lose life they have no claim against the bystander that he not interfere with when the way he interferes involves terminating his life support. So they would have no complaint if the trolley was redirected to them by the bystander. It is true in both this Saving-by-Letting-Die and in the regular Bystander Cases that if the bystander turned to the one, he would provide lifesaving aid to the five and not let them die. However, in the first case he would save five who would have no complaint in being caused to die by the trolley, and in the second case he would save five who would have a complaint in being caused to die by the trolley.

It is true in both cases that if the bystander lets them die, the five would have no complaint against him in particular because he has no duty (and, in particular, no duty to them) to turn the trolley even if it is permissible for him to do so. However, it seems that this is not what is morally decisive if it is permissible for the bystander to turn the trolley in the regular Bystander Case. What seems to be morally significant is simply whether the five have a complaint against the trolley hitting them. If they do, the bystander may refuse to let them be hit. If they do not, he should not turn the trolley toward one person who has a claim not to be interfered with.²⁷

My tentative conclusion about the Bystander Saving-by-Letting-Die Cases is that even they do not show that a mere bystander may not turn the trolley away from killing the five toward killing one in the original Bystander Case. They just show that sometimes when one merely lets die, there will be no people who will have a complaint about what is happening to them, but that will not always be true when one lets die. However, some of these cases do suggest that whether a trolley may be turned could depend on who would turn it; this is because who threatens people with a trolley could affect whether those people have a complaint against being so threatened, a complaint that is needed to justify redirection. For example, the permissibility of turning could vary depending on whether someone who redirects is already involved in saving those who would be threatened by the trolley. On the other hand, who turns the trolley may not matter when it does not affect whether there is a complaint against being hit by the trolley.

Let me conclude this lecture with a final word about what the Trolley Problem is: if both the driver and the bystander are permitted to turn the trolley but neither may topple the fat man or set off the bomb that would kill another person in order to stop the trolley, then the Trolley Problem is about why it is sometimes permissible to kill, even rather than let die, when we come to kill in some ways but not others. So understood, I think the Trolley Problem is a real problem that applies to cases not involving trolleys but with a similar structure, and solving it may require explaining the moral differences among different ways of coming to kill. If this is so, then a detective on the Trolley Problem mysteries cannot (usually) decide whether something wrong was done just by finding out who did it; further investigation would be needed.

NOTES

In 2010 I wrote a short entry on the Trolley Problem for the *International Encyclopedia of Ethics*, edited by H. LaFollette et al. (Malden, MA: Wiley-Blackwell, 2013). The entry was supposed to explain what the problem is, its origins and history, as well as some proposed solutions and other attempts to defuse the problem. In these lectures, I hope to explore some aspects of these topics at greater length.

1. Philippa Foot, "The Problem of Abortion and the Doctrine of Double Effect," *Oxford Review*, November 5, 1967, reprinted in *Killing and Letting Die*, edited by B. Steinbock and A. Norcross, 2nd ed. (New York: Fordham University Press, 1994), 266–79.
2. She said of the driver of the trolley (whom she named Edward) and the transplant surgeon (whom she named David): "Why is it that Edward may turn that trolley to save his five, but David may not cut up his healthy specimen to save his five? I like to call this the trolley problem, in honor of Mrs. Foot's example." Judith Thomson, "Killing, Letting Die, and the Trolley Problem," *Monist* 59, no. 2 (1976): 206. In a later article, "The Trolley Problem," *Yale Law Journal* 94, no. 6 (1985): 1394–1415, she changed her use of the title, applying it to a different set of cases. I shall return to this point below.
3. Judith Thomson, "Turning the Trolley," *Philosophy & Public Affairs* 36, no. 4 (2008): 359–74.
4. See, for example, Judith Thomson, "Physician Assisted Suicide," *Ethics* 109, no. 3 (1999): 497–518; T. M. Scanlon, *Moral Dimensions* (Cambridge, MA: Belknap Press of Harvard University Press, 2008).
5. Thomson, "Turning the Trolley," 369.
6. I first discussed this case in my entry on the Trolley Problem in the *International Encyclopedia of Ethics*.
7. Thomson, "Killing, Letting Die, and the Trolley Problem."
8. Thomson introduced the Bystander at the Switch Case in her "The Trolley Problem," 1397.
9. I discussed cases like this in my "Constraints and You," a paper presented at the American Philosophical Association, Pacific Division in 1983. It became a chapter in my *Morality, Mortality*, vol. 2 (New York: Oxford University Press, 1996).
10. Thomson, "The Trolley Problem," 1399.
11. Thomson, "Turning the Trolley," 363.

12. She shifted her use of the title in her 1985 “The Trolley Problem,” where she first introduced the Bystander at the Switch Case. There she said, “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” (1401).
13. Thomson, “Turning the Trolley,” 368.
14. There are reasons, aside from their inherent interest, for examining her arguments for the reversal. When a major figure whose past work on a problem has been admired and served as a basis for subsequent work by others, one wants to be sure that the reversal is justified, especially since so many still believe there is a Trolley Problem and continue to produce work dealing with it. It is a happy fact that whether Thomson’s reversal is shown to be justified or unjustified, Thomson—either Thomson 1 or Thomson 2—will be vindicated.
15. Thomson, “Turning the Trolley,” 364. This case is like the Trilemma Case presented by Peter Unger in his book *Living High and Letting Die* (New York: Oxford University Press, 1996). (I discussed it in an examination of his book. See “Grouping and the Imposition of Loss,” *Utilitas* 10, no. 3 [1998]: 292–319, reprinted in my *Intricate Ethics* [New York: Oxford University Press, 2007].) Unger’s aims are to show that it is permissible for a bystander to turn from six to three and that if he would do this, he should turn to one instead even if he is the one. He believes that what he calls the Principle of Ethical Integrity requires us to be willing to do to ourselves what we would do to others.
16. Thomson, “Turning the Trolley,” 365, 366.
17. This example was suggested by Larry Temkin. I discussed other examples in my “Supererogation and Obligation,” reprinted as chapter 15 in my *Morality, Mortality*, vol. 2.
18. That is, she first argues for her “Third Principle: A must not kill B to save five if he can instead kill himself to save the five,” and then claims her “Fourth Principle: A may let five die if the only permissible means he has of saving them is killing himself” (“Turning the Trolley,” 365).
19. The permissibility of imposing, and being willing to impose, costs on others that one will not impose on oneself is sometimes referred to as the self/other asymmetry. I discuss it in connection with war, in “Failures of Just War Theory,” reprinted in my *The Moral Target: Aiming at Right Conduct in War and Other Conflicts* (New York: Oxford University Press, 2012), where I consider that it may be permissible to impose costs on enemy civilians that one would not and need not impose on one’s own civilians (or even tolerate being imposed on them by the enemy). I discuss it in connection with permissible collateral harm to civilians, harm that could but will not be absorbed by military personnel, in “Killing in War: Traditional and Non-traditional Views,” also in *The Moral Target*.
20. Thomson, “Turning the Trolley,” 366–67.
21. *Ibid.*, 369.
22. Presumably, this principle is meant to apply only if the five would be killed impermissibly.
23. *Ibid.*, 372.
24. *Ibid.*, 367, 368.
25. See Thomson, “The Trolley Problem,” 1401.
26. I first pointed this out in my *Morality, Mortality*, vol. 2.

27. Notice that “have a complaint” should be interpreted more broadly than “would be treated unjustly by someone.” Suppose it were a natural event, such as a very strong wind, that is moving the trolley toward five people in the original Bystander Trolley Case. If it is permissible for the bystander to turn the trolley in this case, it cannot be because the five have a complaint against anyone in particular that the trolley is headed to them. We will have to allow that they may have a complaint simply because there is no rightful treatment of them taking place. By contrast, rightful treatment of them can be taking place when someone disconnects them from his optionally provided life support.

LECTURE II. HOW WAS THE TROLLEY TURNED?

I

My primary aim in this lecture is to consider some positive proposals¹ that might explain why it is at least thought to be permissible to harm some to save others in two basic trolley cases discussed in lecture 1, one involving a trolley driver as the agent and another a mere bystander. Recall that the basic Trolley Case introduced by Philippa Foot (figure 9) involves a driver in a trolley whose brake does not work when the trolley is headed toward killing five people on the track.² The driver can stop this only by diverting to another track where the trolley will foreseeably kill one other person. For the driver, it is said, it is a choice between killing five and killing one, when all the people are assumed to be alike in morally relevant respects. The basic Bystander Case introduced by Judith Thomson involves the same trolley when the driver is unable to do anything, but a mere bystander can divert the trolley, killing the one other person on the side track. For the bystander it is a choice between letting five die and killing one.

In lecture 1, I considered Thomson's most recent view that whether killing one by turning the trolley is permissible depends on who turns the trolley. She thinks that the driver may permissibly turn the trolley because it is his only way of not committing the greater wrong of unjustly killing five people, but a bystander may not turn it because he will otherwise only let five people die and this is not unjust. I had two major concerns about this view. First, it does not explain why the driver may kill one person by turning the trolley but not by pressing a switch to topple a fat man from a bridge so that he lands in front of the trolley, when his weight would stop (or redirect) the trolley so that it does not kill the five (figure 10). Second, I was not convinced by her arguments that someone who would otherwise only let five die is not permitted to do what someone who would otherwise kill five is permitted to do. However, I was willing to accept that a bystander too may not topple the fat man (in a case like the one in figure 10) and also that only the person who would otherwise kill the five might be obligated to turn the trolley.

At the end of the first lecture, I suggested that if both the driver and the bystander are permitted to turn the trolley but neither may do certain other things that would save the five and kill another person, then the Trolley Problem is about why it is sometimes permissible to kill innocent,

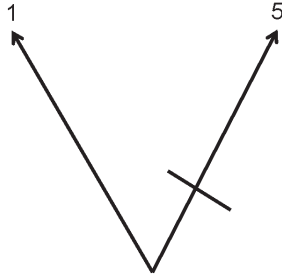


FIGURE 9. Trolley Driver Two-Option Case: Driver can kill five or turn trolley to one.

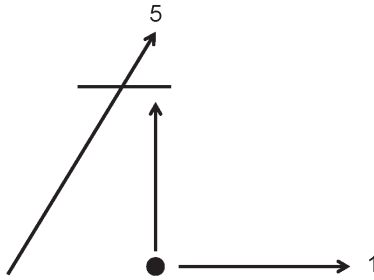


FIGURE 10. Driver Tackle Case: Driver can kill five or topple man in front of trolley, killing him to save five.

nonthreatening people, even rather than let others die, when we come to kill in some ways but not others. In other words, a detective investigating the mysterious Trolley Problem who wants to know if the one person was killed permissibly should (usually) investigate how the person was killed (and also how the five were saved) rather than who did the killing.³ I further said that we could see the Trolley Problem so understood as presenting a challenge to nonconsequentialists who reject the act consequentialist view that we should always do what produces the best consequences. Nonconsequentialists, in some rough way, think that there is ordinarily what is called a side constraint on harming an innocent, nonthreatening person to save five others, even if we assume that five being alive is a greater good and one being dead is a lesser evil. A side constraint is not merely a factor that weighs against producing the greater good; it is supposedly a factor that has priority over producing the greater good. That is, we may produce the greater good only if doing so is consistent with respecting that constraint. One view is that side constraints are individual rights giving rise in others to correlative duties owed to the right bearer and that they

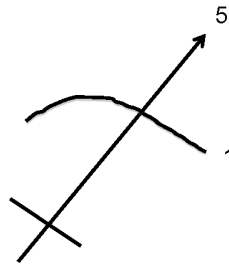


FIGURE 11. Transplant Case: Kill one person for organs to save five.

reflect fundamental aspects of our conception of persons and their status. An example in which we are constrained from killing one to save five is the so-called Transplant Case (as represented in figure 11): if a doctor (or someone else) kills a single healthy person, she will be able to use his organs to save five fatally ill patients (for whose illnesses the single person is not responsible). Yet, intuitively, it is impermissible to kill the one person to save the five. The Transplant Case is presented as a counterexample to act consequentialism. For nonconsequentialists, the challenge is to explain exactly what the side constraint on harming amounts to, what its form is, if it does not exclude turning the trolley and thus killing one person to save five others.

I shall begin this second lecture by considering Thomson's proposal for why many people believe, mistakenly she thinks, that the bystander may turn the trolley, since her explanation focuses on the "how" question. I shall then consider an alternative proposal for why both the driver and the mere bystander may turn the trolley that focuses on different ways in which the trolley can come to be turned or otherwise stopped from killing the five. These different ways amount to both different ways in which harm to some person(s) comes about and also different ways in which the good of other persons not being harmed comes about. This proposal claims that if harm to some and prevention of harm to others come about in certain ways, it will seem (and perhaps actually be) permissible to turn the trolley. The proposal also attempts to characterize in a general way when it is and is not permissible to harm some to save others.

Though I will present some criticisms of this proposal, I will nevertheless consider how something like it relates to the moral distinction between killing and letting die and then whether all elements of the proposal are required, given what others who would also kill rather than let die seem to be permitted to do. This will lead us to consider the role of partiality

and impartiality in the decision to do what harms others. In conclusion, I will consider some cases that again raise the question of what the Trolley Problem is.

I should preface what is to come with a warning. Typically, people who work on the Trolley Problem employ the following methodology: They start off with some basic cases and suggest a principle to explain and perhaps justify intuitive judgments in those cases. But then other cases are imagined that show the proposed principle to be wrong or in need of revision. Perhaps a new principle is then suggested to explain and justify the judgments in the original cases plus the new ones. Then other cases may be raised that make problems for the new principle, and on and on. If a principle seems satisfactory in terms of cohering with intuitive judgments about cases, then the issue is to see if the principle reflects some morally significant underlying concepts or values.⁴ In considering suggestions for a principle here, I will not be attempting to deal with all the many cases that have to be correctly accounted for. So any principle that seems to shed light on the basic cases may have to be, at the very least, modified. Here I just hope to show some directions that have been taken and in which we might go. So while it may seem like we are considering many hypothetical cases, I assure you they will not be enough!⁵

II

Thomson concludes her 2008 article “Turning the Trolley” by offering an error theory for why it has been mistakenly thought that a mere bystander who would otherwise only let five die is permitted to turn the trolley and kill one person.⁶ She believes that many people’s judgments about the abhorrence (and hence permissibility) of killing vary with how “drastic an assault on the one the agent has to make.” This leads them to think, in her view mistakenly, that the bystander causing death by pushing the fat man is impermissible, but the bystander causing death by turning a trolley is permissible, because in the latter case, “if he proceeds, he will bring about that more live by merely turning a trolley.”⁷ But if she was right about this, people should also think it is permissible to turn a trolley in the following Two-Trolleys Case (in figure 12) where the bystander is Mr. X: Suppose one trolley is headed to the five. We can redirect this trolley away from the five if we just turn a second trolley, running on a track where it would harm no one, into the first trolley. However, on its way to redirect the first trolley, the second trolley will kill another person as a mere side effect. I doubt that people would think that merely turning a trolley makes turning the

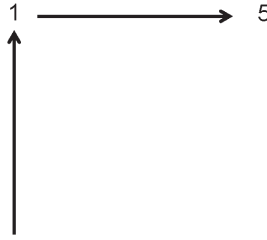


FIGURE 12. Two-Trolleys Case: Bystander (X) can let five die or turn second trolley to stop first; second trolley will kill one.

second trolley permissible. And if she was right, should not people also think it is permissible for someone to topple the fat man in front of the trolley not by pushing him over, but by merely turning a second trolley that goes up on the bridge, toppling him? I doubt that people would think that only turning a trolley in this case will make toppling permissible. Hence, an explanation of why people think turning the trolley in the standard Bystander Case is morally permissible will have to point to some factor other than that it involves “just turning a trolley.”

III

Consider some alternative proposals to explain why people think (perhaps justifiably) that turning the trolley in both the Bystander and the Driver Cases is permissible. (Note that people may not be able to articulate these proposals, which nevertheless underlie their judgments.) One proposal is that if the one person is hit, he would stand in exactly the same relation to the same trolley threat that the five would if they were hit. They would all come to be threatened by the same trolley going to them, by its going either to the five or to the one. This contrasts with what happens in the Two-Trolleys Case where the trolley that threatens the five is not the trolley that would threaten the one. (This is also true in the case discussed in lecture 1 in which someone sets a bomb to move the trolley from the five and the bomb kills the one other person, as represented in figure 13.) Toppling a fat man from a bridge in front of the trolley so that his body stops (or redirects) it also involves something different happening to him from what would have happened to the five (that is, they would have the trolley sent to them, whereas he would be sent to the trolley). Is it perhaps the potential victims standing in the same relation to the same threat that distinguishes the Bystander and Driver Trolley Cases morally from other cases of killing to save others?

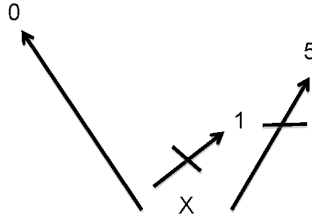


FIGURE 13. Bomb Trolley Driver Case: Driver can kill five or set off bomb to stop trolley; bomb kills one.

I think that one problem with this proposal is that it is too narrow, in the sense that these trolley cases may be one type of a more general class of cases in which the permissibility of harming is explicable in essentially the same way. We might come to see this by considering another aspect of turning the trolley in the basic cases: the good of five being saved seems to be the mere noncausal flip side of removing what threatens them (the trolley). That is, in the circumstances of these cases, where there are no other threats to them, the five being saved is simply the trolley moving away. The relation between the trolley moving away and the five being saved is not a causal relation but rather seems to be a constitutive relation.

By contrast, in the Two-Trolleys Case, the second trolley, which is not a threat to the five, would cause the first trolley's moving away from the five, and the first trolley's moving away would have the noncausal, constitutive relation to the five being saved. In the basic Trolley Case, the means (namely, turning the trolley) that has the constitutive relation to the five being free of the threat and to their being saved is what will directly cause the lesser harm to the one person on the other track, since the trolley that is moving away will hit the one. This way in which the means of saving the five comes to harm some and help others is causally different from the way in which the second trolley in the Two-Trolleys Case would come to harm some and help others: that second trolley, whose movement does not have the noncausal, constitutive relation to the five being free of a threat and being saved, both causes the first trolley to move away and also directly harms one other person. Perhaps it is impermissible to use means that have such a mere causal rather than a noncausal or constitutive relation to removing a threat and saving the five when those means will also directly harm someone else. When the fat man is toppled so that he stops the trolley on its original route to the five, his being in harm's way is, and is brought about by, a mere causal means to producing the greater good of five saved. So the harm to him is also not produced by means that have a

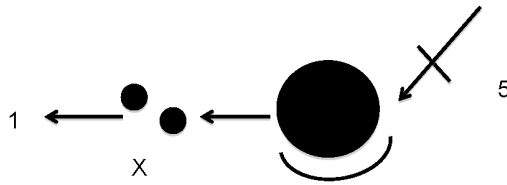


FIGURE 14. Lazy Susan Case: Bystander (X) can let five die or turn lazy Susan; five in safe area cause rock slide that kills one.

noncausal relation to stopping a threat and to the greater good; rather, his being hit causes the threat to stop and thereby causes the greater good.

The noncausal, constitutive relation of the trolley moving away to saving the five in the Trolley Case makes the case very close to one in which the five being saved itself leads to harm to another person. For example, suppose that we cannot redirect the trolley away from them, but we could move them away from the trolley by turning a swivel table—a large lazy Susan—on which they are seated. However, in Lazy Susan Case I, their being in the safe location also involves another person, near the side of the lazy Susan, being pushed into the trolley. In Lazy Susan Case II, their being in the safe location causes rocks in the area to fall, killing another person (as in figure 14). In these cases, the fact that five are being moved away from the trolley has a noncausal, constitutive relation to their being safe from the trolley that threatened them and, in the context where this is the only threat, to their being saved. In Lazy Susan Case II, their being away causes another entity, rocks independently in the environment, to become a threat to someone else. (In the basic Trolley Case, by contrast, moving away an entity independently in the environment [that is, the trolley], thereby causing it to become a threat to someone else, has a noncausal, constitutive relation to the five being safe from that entity. This is possible because the trolley threatens the five originally, unlike the rocks in the Lazy Susan Case that do not threaten the five.)

Suppose we think that moving the five away from the trolley in the Lazy Susan Cases is permissible for essentially the same reason that turning the trolley from the five in the Trolley Case is permissible. Then it would not be crucial to the permissibility of turning the trolley that we are not pushing another person into the trolley. Nor would it be crucial that we are redistributing the same threat that already faced the five to another and that the other will stand in the same relation to the same threat that the five did. After all, in Lazy Susan Case II, it is rocks that are moved by the five being safe that threaten the one person, not the trolley that was headed to the five.

In all these cases in which I have said that turning the trolley has as its noncausal flip side the five being free of that threat and being saved, it is strictly true that the flip side is only the first instant of their being unthreatened.⁸ This is a component of whatever period of life is sufficient to make it the case that their being alive is a greater good relative to the death of one other person. (This component also helps cause the future instants of life that make up the future life of these people. However, it is not a mere cause of the future life; it is also a component of that life.) We should not turn the trolley killing one person if the five would live for only one instant, but only if we have good reason to believe they will go on living long enough. Similarly, in cases where I have said that five being saved causes the death of the one person, it is strictly only the first instant of their being free of what threatened them that causes the one's death, as when the lazy Susan lands them where their weight causes rocks to fall on another person. This first instant of their being safe is a component, as well as a cause, of what we should reasonably expect will be a period of life for the five people long enough to be a greater good relative to the lesser evil of one person dying.

An important aspect of the discussion so far bears mentioning. Foot drew her major contrast between the impermissibility of the doctor killing in *Transplant* and the permissibility of the driver killing in *Trolley*. Similarly, most have emphasized the distinction between turning the trolley and pushing the fat man to stop the trolley, which is (as Thomson pointed out) the version of the *Trolley Case* that is like *Transplant* with respect to how the one will be treated. However, in my own past work on the *Trolley Problem*, I emphasized a different contrast based on another of Foot's cases (which I mentioned in lecture 1), the *Gas Case*.⁹ In arguing against the Doctrine of Double Effect (DDE) as an explanation of why the trolley may be turned, she noted that a doctor may not use a gas needed in surgery to save five people when it will foreseeably cause a bystander to die. This is so even though his death (or mere involvement) is not intended or causally useful to save the five, unlike getting organs in *Transplant*. I argued that given what was true in the *Gas Case*, it was crucial to explain why turning the trolley is permissible even though it too is a means to saving the five that we foresee will kill one person whose being hit is not intended or causally useful to save the five.

One problem with Foot's discussion may be that while she considered cases in which (1) a person's being harmed is intended and is a means to helping others (*Transplant*) and (2) harm is a side effect of a mere causal means to helping others (*Gas*), she did not consider (3) cases in which harm

is a side effect of others already being saved (Lazy Susan) and (4) cases in which harm is a side effect of means that have a noncausal, constitutive relation to the good they produce (Trolley). It is the latter two types of cases I focused on. She also did not recognize that mere causal means to saving the five (like the gas) might be used if it killed someone indirectly by, in particular, affecting some entity independently in the environment that would kill another person. For example, suppose the gas we introduce that is needed to help the five in Foot's case is itself harmless to everyone. However, spraying it changes air currents so that some fatal germs that had previously been closeted move to one person, killing him. Even if we knew this would happen, it seems permissible to use the gas to save five. Similarly, suppose the second trolley sent in to move the first one did not kill a bystander. However, its movement caused rocks independently in the environment to fall, killing someone. Then perhaps it would be permissible to send in that second trolley to save the five.¹⁰

We could summarize some of these views about permissibly and impermissibly seriously harming innocent, nonthreatening people to save others as follows: Actions are permissible if the greater good or a component of it (or means having these as a noncausal flip side) leads to lesser harm even directly. Actions are impermissible if mere means that produce the greater good (like the bomb or second trolley) cause lesser harm at least directly, and actions are impermissible if mere means cause lesser harms (such as toppling people in front of a trolley) that are mere means to producing greater goods. (That an act is permissible does not imply that those who would be harmed by it may not also permissibly resist it and its effects.)¹¹ This is a rough description of one version of what I have called a Principle of Permissible Harm (PPH).¹² It might be described as a "downstreamish" principle in that it implies that the lesser, direct harms may permissibly be causally downstream of the greater good, components of it, or means having these as their noncausal flip side.

I say this PPH is a "downstreamish" view rather than a "downstream" view because in certain other cases, removing the trolley threat that would kill the five people will result in harm to the one person that itself plays a further necessary causal role "upstream" to the greater good, either sustaining a component of the greater good by deflecting threats to it or helping to produce the greater good. The most famous example of this sort is the brilliant Loop Case introduced by Judith Thomson.¹³ In this case, if the trolley is redirected, it would loop back to kill the five were it not that its hitting the one person on the track stops it. Here one person being

hit is a necessary causal means to the five being saved as much as in the Topple Case. Yet the PPH might imply that turning the trolley is permissible because, unlike what is true in the Topple Case, the one's being hit is caused by the turning of the trolley threat that has a component of the greater good as its flip side. That is, we turn the trolley because this is a way to stop the trolley from killing the five in its initial direction. If we just consider what this leaves us with, abstracting from any new threats created by what we do in redirecting, we see (what I have called) the "structural equivalent" of the greater good, that is, the five free of all threats they faced independently of our redirecting. This is a particular type of component of the greater good that would be the greater good if only it were not undermined by the new threat of the looping trolley created by our redirection. We turn the trolley because (or on condition that) it will hit the one person, for this will prevent the component from being undermined; the component will be sustained and become the greater good. But according to the PPH, we should not do anything (like giving the trolley an extra push not needed to get it away from the five) in order that it hit the one, for this would do something that did not have any component of the greater good as its noncausal flip side. The extra push would be a mere means to the one being hit, like toppling the man in the Topple Case.¹⁴

This discussion of the Loop Case is intended to show that, according to the PPH, it is important how we bring about the harm that will have a necessary causal role in saving the five. But it also shows that it is important what the harm is causally necessary to bring about. To make this point clearer, consider what I call the Tractor Case:¹⁵ The five toward whom the trolley is headed also have a deadly tractor headed toward them. If we turn the trolley away, it will hit one person whose being hit will stop the tractor. Is it permissible to turn the trolley in the Tractor Case as it is in the Loop Case? In the Loop Case the threat to the five of the trolley coming back to hit them is itself produced by the redirection, and so abstracting from this new threat, we have the structural equivalent of the greater good as a flip side of redirection. It needs to be sustained and will be sustained by the hit produced by the same redirection. In the Tractor Case the one person's being hit interferes with a threat to the five that exists independently of what we do to save them. So when we turn away the trolley and abstract from the further effects of what we do, it seems that we do not yet have any component of the greater good as a noncausal flip side, for the five are still under a fatal tractor threat. The fact that the tractor threat exists independently of what we do to remove the initial threat of the trolley means

that we need to produce another outcome—the removal of the tractor—when we redirect, rather than just sustain the removal of the trolley that we have already produced, and the harm to the person needed to produce this outcome is not caused by means that already have a component of the greater good as a noncausal flip side. So this productive harm is not downstream from a component of the good. The possible difference in the permissibility of redirecting in the Loop and Tractor Cases suggests that in considering the permissibility of actions that will harm people, we should consider both how the lesser harm comes about and what the lesser harm is needed to bring about (and that these can be related).

The PPH is one version of an approach that focuses on different ways of bringing about goods and harms and different causal or noncausal, constitutive relations between goods, harms, and means of producing them. An additional possible hypothesis is that supervening upon these differences are different relations between the people harmed and those saved from harm. In particular, one suggestion is that the permissible ways to harm some to save others involve some people being substituted for others (when this is done in certain ways) with regard to being harmed, whereas the impermissible ways to harm involve some people being subordinated to others, either by coming in certain ways to be used for others (as in toppling someone to stop the trolley in the Topple Case) or by direct harm to them being considered less important relative to the use of certain means (such as sending in the second trolley) when these means are causally useful for others. Harming by substitution could be permissible even if harming by subordination was impermissible. Ultimately, identifying certain relations between people as permissible and impermissible could reflect a certain conception of persons and their status. It is not the point of this lecture to investigate these deeper possible meanings of the PPH I have described, but it is useful to keep in mind that finding the deeper meaning of any PPH (this or another) is both important to a complete understanding and justification of action in the trolley cases and to showing how the Trolley Problem is connected to our conception of persons and their status.

There are also problems with this proposed PPH, I believe.¹⁶ Like earlier proposals, the problems stem from generalizing from too narrow a set of cases. Turning the trolley seems to be permissible in cases that do not satisfy the conditions of the PPH. Consider Tractor Case II, which is like the original Tractor Case except that the person's being hit on the side track has no causal role in stopping the tractor. Rather, the tractor is stopped by a switch that is pressed by the trolley as it is turned away from

the five. In this case, no more than in the original Tractor Case does the moving away of the trolley have the greater good or a component of it as its noncausal flip side, since the five are still under the threat of a deadly tractor before the switch is pressed. Yet it seems permissible to turn the trolley away even though it kills one person, given that the tractor threat will be taken care of by innocent means of the trolley pressing a switch. In this case, it seems correct to say that the turning trolley has a causal, not a noncausal, constitutive relation to the greater good or a component of it. More generally, there may be many cases in which turning the trolley is permissible, yet it removes only one threat that would impede the greater good. Other innocent means, even independent of turning the trolley, are needed to deal with other threats and thus produce the greater good. (Similarly, there may be many cases in which the five moving away from the trolley [as in the Lazy Susan Cases] does not have a component of the greater good as its flip side, since the five are still subject to another fatal threat, yet their moving away is permissible given that innocent means will deal with the other threats.) The PPH as described draws too close a connection between turning away a threat that is constitutive of the five not being subject to that threat in its original form and turning away a threat having a noncausal, constitutive relation to the greater good or some component of it.¹⁷

Nevertheless, there is still a distinction between the one person being hit by a mere means to removing a threat, such as the second trolley, and the one person being hit by the removal of a threat to the five (or the removal of the five from a threat) that would impede the greater good. Turning the trolley in Tractor Case II may have a causal relation to the greater good, but it also has a noncausal, constitutive relation to the five being free of something that would impede the greater good, and so it is not a *mere* means to the greater good like the second trolley.

In sum, the two Tractor Cases and the Loop Case suggest that explaining why it is permissible to turn the trolley may rely on its being a threat that is removed that kills one person and how removal helps bring about the greater good (for example, as in Tractor Case II and in the Loop Case versus in the original Tractor Case). It seems that when the one person being hit by redirection is a causal means to producing a greater good, but not when the harm caused by redirection is a mere side effect or needed to sustain what we have produced, the harm should have been caused by means that have a noncausal, constitutive relation to a component of the greater good and so be downstream from it.

Another possible problem with the proposal for a PPH is related to the Saving-by-Letting-Die Case (from lecture 1) in which someone would have to turn a trolley toward interfering with ongoing life support that he is optionally providing to the five people rather than toward one other person to whom he is not providing such aid. It seems to me that the greater good is the five surviving, yet it was suggested in lecture 1 that either the driver or a bystander should terminate his aid with the trolley rather than redirect to kill someone not receiving such aid. If this is correct, then the PPH proposed is too broad. I suggested that in this case the five would have no right not to have the person providing their life support terminate aid to them, and so the five would have no complaint if he directed the trolley to end the life support he was providing. By contrast, the one person would have a complaint in being hit. Hence, it is important that the greater good be a state whose not coming about someone would have a complaint against; not just any greater good will override the complaint of the person to whom the trolley is redirected. In the basic trolley case, the five would have a complaint if they were hit by the trolley driver, and this is one reason the bystander and the driver may help them.

IV

Despite the problems with this proposed PPH, for purposes of illustration let us consider how such a proposal (suitably modified) that focuses on relations between harms and goods and between some potential victims connects with the question of whether a driver who would otherwise kill the five may turn the trolley, but a mere bystander who would let them die may not. Doing this may help illustrate a way nonconsequentialists could insist on the moral distinction between killing and letting die without also claiming that the distinction helps draw a line between what the driver and bystander are permitted to do (at least in many cases).

A

First consider what I will call the “Agent-Victim Killing–Letting Die Distinction.” Some nonconsequentialists think that we as agents must make greater efforts to avoid killing someone than to avoid merely letting someone die, because of differences in what we would do to the victim. I have argued that this is true at least when killing involves depriving someone of life that he would continue to have at that time independently of the agent’s provision (for example, he would go on living if the agent and his life support devices did not exist, holding all else constant). When we

let die, by contrast, someone loses life he would not continue to have at that time independently of the agent's provision (for example, he would die if the agent and his life support devices did not exist, holding all else constant).

Some of those who focus on the Agent-Victim Killing–Letting Die Distinction are saying that not only is there a moral difference in what an agent does to someone in killing him or letting him die, but what an agent may permissibly do, and so what relation he may have, to other people in order to avoid the relation of killing someone is different from what he may do, and so what relation he may have, to other people in order to avoid the relation of letting someone die.

However, there is a second type of relation, which holds *among the people whose lives are at stake*, rather than between those people and the agents who would affect them, and a question is whether the relation between these people is also sensitive to the distinction between killing and letting die. I shall call this the “Intervictim Killing–Letting Die Distinction,” where the victims I have in mind are those who were originally threatened and those who would die if we saved those originally threatened. What relations hold between these people may depend on what an agent does to the people. But it is possible that at least certain relations *between the people* whose lives are at stake will be the same whether those relations come about because of what is done by an agent like the trolley driver, who would kill some rather than others, or by an agent like the mere bystander, who would kill some rather than let others die. For example, suppose the driver of the out-of-control trolley is, as Foot and Thomson see it, in the process of killing the five. In this case, he is in the process of making them worse off, depriving them of lives they would have had. If he redirects the trolley, he will be making someone else worse off in the same way. This is the Agent-Victim description. Put in Intervictim terms, the one person will be killed and, therefore, be made worse off as the alternative to others being killed and made worse off. This Intervictim description could be true whether it is the driver who turns the trolley or a mere bystander who turns it. Even though as an agent the mere bystander would let the five die and so not benefit them rather than make them worse off, the five would still be killed by the trolley and be made worse off.

Suppose, by contrast, that we thought of the five as already worse off than they had been in virtue of the threat to them, so that in turning the trolley, the driver would be improving their condition, albeit by doing what makes one other person worse off. Considered from the Intervictim

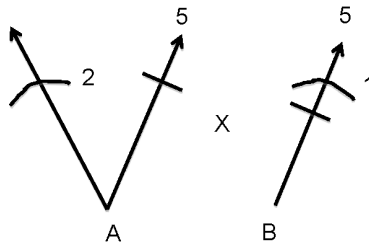


FIGURE 15. Intervictim/Agent-Victim Case: Bystander (X) can (a) let five die or (b) turn trolley that saves five and topples two or (c) topple one to stop trolley and save five.

perspective, the five would be improved by the trolley being turned away, and that turning makes the one person worse off. This could be true as well if the mere bystander turned the trolley. The fact that the driver bears a different agent relationship to the five than the mere bystander does need not mean that in these respects the individuals whose lives are at stake have different relations to each other depending on which agent acts. The same descriptions could be true of them whichever agent acts.

A further example may help to reinforce the distinction between the agent's relation to a state of affairs (at least put in terms of killing or letting die) and the Intervictim character of the state of affairs. Suppose a trolley driver faces a choice among (a) killing the five, (b) turning the trolley away from the five onto a track that runs on a bridge so that two other people get toppled from the bridge, or (c) pressing a switch that moves one fat man off a bridge so that he stops the trolley headed to the same five people. These choices are represented in figure 15. Those who emphasize the killing–letting die distinction may see the possible Agent-Victim relations here as killing five or killing two or killing one, in which case the one should be killed. However, the Intervictim relation between those whose lives are at stake will be very different depending on what is done. In option (a), two people will be killed as a consequence of removing a threat to five. In option (c), one person will be killed as a result of toppling him to stop the trolley going to the five. These different Intervictim relations are a function of how the people come to be killed and their role in producing the greater good. In one case, people are killed as a consequence of the removal of the threat to others; in the second, one will be killed as a consequence of a mere means (such as a switch) used to move someone so that he is a means to remove the threat to others. This could make it permissible to kill the two people but not permissible even for the driver to kill the one other

person. We must consider the relations between the people whose lives are at stake that would come about as a consequence of what is done. It could be that the same Intervictim facts that make some options permissible for the driver also make them permissible for the bystander.

The rough proposal for a PPH that I described earlier identifies such relations between potential victims. For example, in some cases the person killed dies as a result of the greater number being saved (as when we move the five on the lazy Susan, and their being in a safe area causes rocks to kill someone else). Here a component of what can reasonably be expected to be the greater good leads to lesser harm, and then, the proposal suggests, the Intervictim relation is substitution with respect to being threatened and is permissible. In the Two-Trolleys Case, the person is killed by a mere means to what would causally produce the greater good, and then his relation to those saved, the proposal suggests, is subordinating and not permissible.

It may be that the driver is permitted to kill fewer people by turning the trolley only because the Intervictim relation that would be created between the potential victims makes it permissible for him to act. And if the Intervictim relations are the same when the mere bystander turns the trolley, it may be as permissible for him to act even if, unlike the driver, he has no duty to act because as an agent he would only be letting five die. Suppose there is something about the Intervictim relation between the one and the five that makes it impermissible for the mere bystander to bring it about. When the one person should not be made to stand in that relation to the five just so that the bystander can avoid the relation to the five of letting them die, it is not clear why the one may be made to stand in that relation to the five just so that the driver can avoid the relation to the five of killing them.

Of course, if the relation between potential victims would be different depending on what an agent does, one could also say that the Agent-Victim relation is different in virtue of this. Namely, if the agent does one thing, the agent makes a victim stand in one relation to another potential victim; if the agent does another thing, he makes a victim stand in a different relation to another potential victim. However, this further description of the Agent-Victim relation is dependent in the first instance on the potential victims' relations to each other, and it is the Intervictim relations that are significant for the permissibility or impermissibility of the Agent-Victim relations.

B

Suppose the relations between the potential victims determined whether it is permissible or not to kill some to save others. How would this affect

the nonconsequentialist claim that there is a stronger duty not to kill than not to let die? It would not affect one implication of this claim, namely, that those who would kill the five have a duty, and should pay great personal costs, to not kill them, while those who would only let the five die (including let them be killed) may have no duty, and need not pay great personal costs, to save them. Furthermore, it need not affect the claim that there are side constraints, such as its being impermissible to topple the one person in the Topple Case. However, if relations between victims determined the permissibility of killing them, then the duty not to kill would not, in general, take precedence over a duty (should it exist) or desire not to let more people die.

This would be one instance in which what is, in general, stringent by the measure of there being a duty to do it or having to pay a lot to do it need not take precedence over what is weaker by those same measures. Here is another such case. I may be morally required to pay a lot of money to keep a contractual promise in business but not similarly required to pay as much to save the life of someone drowning. Yet if I have a choice between doing an optional, even costly, rescue of the drowning person or else fulfilling the contractual duty, it is permissible for me to do the rescue.¹⁸

v

Needless to say, many may not accept the proposal for a PPH, even amended, and its relation to the trolley cases and to the question of killing versus letting die as I have described them. Hence, I wish to pursue yet another approach to establishing the relation between how a killing occurs and the permissibility of killing, whether the alternative is killing others or letting them die. It begins by considering some other people besides the mere bystander who would also only let die if they did not kill. The aim is to see how much is thought to be permissible for these people and how the line between what is thought to be permissible and impermissible might be drawn.

Among those who would only let die if they did not do what kills another are the five dying of organ failure in the Transplant Case and the five to whom the trolley is headed in the Trolley Case. If the five were able to but did not kill someone else to save themselves, they would let themselves die. Let us begin with the five in Transplant. Even though their own lives are at stake, I believe it would be impermissible for them to kill one person for the organs they need in the way imagined to be necessary in the Transplant Case. I believe it would also be impermissible for them to

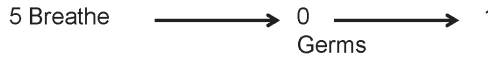


FIGURE 16. Five Breathe/Germs Case: Five let themselves die or five breathe and move germs in area, killing one.

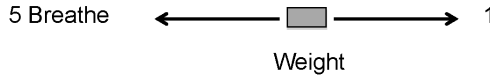


FIGURE 17. Five Breathe/Weight Case: Five let themselves die or five breathe and move weight from them, killing one.

release a gas that they need to cure themselves if the gas will as a side effect kill someone else.

However, suppose that the five could successfully treat their medical problem by means that harm no one. Then they will begin to breathe normally, and their expanding chests will move some fatal germs in the atmosphere that would otherwise have been safely at rest. The germs will reach one other person and kill him (as in figure 16). Is it impermissible for the five to successfully treat themselves when they know all this will happen? Must they let themselves die instead? I think it is permissible for them to bring about their normal state even though their normal breathing will then turn an entity in the environment (the germs) into a fatal threat to others. I emphasize that this permissibility is not merely a matter of their being excused in acting wrongfully because they are in extreme circumstances; rather, they do not act wrongfully at all. Notice that in this case, unlike in the Trolley Case, the five are faced with one threat (organ failure), and their breathing normally causes another threat (that is, moving germs). Even though they create a new threat to others rather than redirect an existing threat from themselves, I think their action is permissible. In addition, as I noted earlier, if they need to release a gas to cure themselves and this gas in itself is harmless to anyone, it seems they may use it even if the gas being in the atmosphere changes air currents and so fatal germs, independently in the environment, move to kill one person.

Now suppose that what interferes with the organs of the five and their normal breathing is a heavy weight that is on them. If by some great effort they could breathe normally just once, their chest expansion would push away this heavy weight. However, it would then roll on to kill one other person (as represented in figure 17). Once rid of the weight, the five's organs and breathing would recover. (Their breathing normally, which

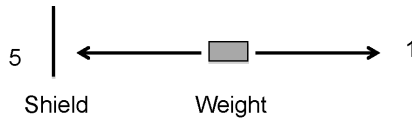


FIGURE 18. Five Shield/Weight Case: Five let themselves die or weight moves off of five's shield to kill one.

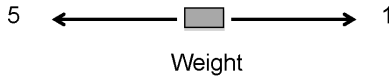


FIGURE 19. Five Move Weight Case: Five let themselves die or move weight away, killing one.

moves the weight, is not the first result of their having recovered from a direct threat, as in the previous case. However, it is a component of their being alive for a longer time.) In this case, we have an entity present in the environment, independently of the actions of the five, that will either threaten them or harm the one. I think it would be permissible for the five to breathe normally in this case, thus removing the threat from themselves, though it will kill another.

What if the five know that once the weight is upon them there will be nothing they can do to save themselves? So they set up a defensive shield against it (perhaps only shielding themselves with their hands). They do this knowing that the approaching weight will be repelled by the shield and kill one other person instead (as represented in figure 18). A shield is a device that does not itself lead to harm to others independent of something else interacting with it (by contrast to a dangerous gas). For moral purposes, the five using a shield seems close to maintaining their normal state. It would produce harm to others only if an entity independently in the environment interacts with what maintains their normal state, similar to when their normal breathing would itself push an oncoming weight away. It seems to me that the five shielding themselves is permissible despite its harmful effect on another person.

If the five may move the weight by breathing normally or by shielding themselves from it, the question that remains is whether they may simply move the weight away—for example, with their hands—in order to breathe normally and not die, though the weight will then kill one other person. (See figure 19.)

VI

In the Trolley Cases, the same issues just described in variants of Transplant could arise for the five if they do not let themselves die, though this will result in the death of one other person. (Hence, some of the diagrams used for the variants on Transplant can do double duty for the cases I will now discuss.) Even though their lives are at stake, the five toward whom the trolley goes are not, I believe, permitted to use a device to topple the fat man in front of the trolley, nor may they use a bomb that will stop the trolley but kill someone else. However, suppose the trolley is coming down a track toward one person, the five are between it and him, and their being hit will save him. It is permissible for them to run away or duck even if the one person is killed instead, since they are thus terminating protection that they are not required to provide him (as in figure 20).

What if, in another case, the five are alone on the track and the trolley is preprogrammed so that if they move away from it, it will start up in another direction, hitting one other person? They are permitted to move away, I believe. Alternatively, what if the five move away from the trolley and wind up in a safe area and their being in this area causes some rocks to fall that kill another person (as in figure 21)? I believe it would be permissible for them to escape the trolley in this case too. And if their merely breathing normally caused the trolley to move away from them toward another, I believe they may breathe. These cases fall into the class of the first part of the good of the five being alive causing an entity—either the trolley that threatened them or rocks—to become a threat to another person.

If the trolley is headed only to the five, they may also, I think, shield themselves from it, even if its interacting with the shield will cause the trolley to be deflected toward another person. Here it is not a component of the greater good itself that leads to the lesser harm but a device that

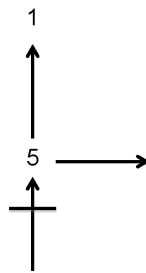


FIGURE 20. Five Duck Trolley Case: Five let themselves die or move away, and one is hit by trolley.

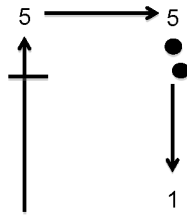


FIGURE 21. Five Move to Safety Case: Five let themselves die or move to a safe area, causing rock slide that kills one.

maintains the good, without causing harm to others independently of the trolley interacting with it.

The question remains whether those who can do all these things rather than let themselves die, even when what they do kills one other person, may simply move the trolley away from themselves so that they are safe. This is like the question, asked earlier, of whether the five on whose organs a weight is pressed may just remove the weight when it will then roll onto someone else.

What could morally distinguish what it is agreed the five may permissibly do from turning the trolley? Here is a proposal: If they push the threat away (or press a switch that does so), the five would get rid of the threat by actively and deliberately doing something to it, just as they would do in setting a bomb that kills someone. My fundamental concern with this proposal is that it draws no moral distinction between actively and deliberately doing something to an entity that is a threat to the five and to one that is not. The former can constitute the five being free of a threat; the latter does not. The alternative view is that the five removing a threat to themselves is in itself as permissible as their breathing normally or moving to safety when this results in their being free of a threat. If the five breathing normally and being in a safe area are permissible acts though they cause an entity in the environment to harm someone else, then it seems their moving the trolley threat away should be permissible even when this causes it, an entity in the environment, to be a threat to someone else.

VII

Before we consider what implications, if any, cases where the threatened people need not let themselves die may have for mere bystander (or driver) cases, let us consider some other potential victims who face a choice between letting themselves die or killing others.

Suppose *one* person has multiple organ failure that is causing his death. He is not permitted to simply kill five people to get their organs or to use a gas to save himself that will kill five. But suppose the one person was able to cure his organ failure by innocent means so that he could breathe and his chest expanded properly. Does the fact that this would move germs in the atmosphere to *five* other people mean that it is not permissible for him to cure himself? Intuitively, I think not. Suppose a weight pressing on him is causing his problems. If he could breathe normally once and this moved the weight away, is he permitted to do this if the weight will roll onto five others, killing them? I think he is. (Illustrations for these cases and others to follow just involve switching the location of the numbers 1 and 5 in figures 17–22 used earlier.) These conclusions would be consistent with the nonconsequentialist view that sometimes it is permissible to do acts that do not produce the best consequences; it all depends on the nature of the acts.

Suppose the trolley is originally headed to *one* person and the question is what is permissible for him to do to save himself even though it results in five people dying instead. He may not, I believe, topple five fat men in front of the trolley to stop its hitting him, and he may not use a bomb to stop the trolley when it will kill five people as a side effect. However, there seem to be other things he would be permitted to do. Suppose the trolley is headed to him as he stands on the track in front of the five people. It would be permissible—not merely wrong and excusable—for him to duck or run away, knowing the trolley would then go on to kill the five. From a non-consequentialist perspective, he has no obligation to stay merely to provide protection that will produce a greater good. When I originally introduced the nonconsequentialist view, I contrasted it with act consequentialism on the issue of side constraints that prohibit us from always doing what can produce the greater good. But nonconsequentialists are also commonly contrasted with act consequentialists on the issue of persons having no general duty to promote the greater good at great personal cost, even when doing so would violate no side constraint. Hence, someone is not always morally required to act for the impartial great good but may have what is known as a personal prerogative to act from his partial point of view.

It also seems to me that the one person is permitted to shield himself from the trolley even though the trolley hitting the shield will be deflected toward five people. May he move to safe ground when he knows that his being there will cause rocks independently in the environment to fall on and kill five people? I believe it is permissible for him to move.

So far, surprisingly, it seems that the one person may permissibly do a lot of what the five are permitted to do rather than let himself die. This is so even though his actions cannot be permissible because greater good or a component of it (or means having these as its flip side) may cause lesser harm, since his being saved is a *lesser* good and the five being harmed is a greater harm. Of course, there must still be some important good at stake for the one to justify his doing various things that will foreseeably harm people. But the focus here is totally on the type of action he takes—how he brings about a good, how he harms, and how the good and harms are related—rather than on what is at stake, namely, greater good coming about at the cost of lesser harm. In addition, the question arises for the one as for the five whether, given all he is permitted to do, he may simply push the trolley away from himself when it will then go toward the five.

This discussion of the single person suggests that any possible PPH should be seen as having two parts: a part that distinguishes different ways of bringing about harm to innocents and another part that applies these distinctions when it is a question of greater good being produced at the cost of lesser harms. Its first part would seem to still govern the behavior of the single individual even when its second part does not apply.

VIII

Are there implications from what the five and the one to whom the trolley is originally headed are permitted to do for themselves for what a mere bystander is permitted to do? Why should we think there are any implications? Because in thinking about what it is permissible for the one or the five to do, we have not relied on any view that implies that people who are themselves threatened may do just anything they need to do to save themselves. Hence, our conclusions about the one and the five may generalize to the bystander who would kill rather than let die. However, suppose the bystander was permitted to do for the one or the five what it is agreed they could permissibly do to save themselves, and suppose also, for argument's sake, that the five and the one were permitted to turn the trolley from themselves though it would kill some innocent(s). Then it would turn out that the mere bystander might be permitted not only to turn the trolley from five to one but to turn it from one to five. This seems like a rather shocking conclusion to reach.

But it need not be so shocking if the bystander is a “mere bystander” only in the sense that he will not kill anyone if he does not redirect the trolley. He might still be a close friend of the one person or his bodyguard.

Then it might not be surprising that the bystander is permitted either to help the one person act permissibly or to do in his stead what it would be permissible for the one person to do. From a nonconsequentialist perspective, a bystander is not always morally required to act impartially but may have a personal prerogative to act from a partial point of view.

However, what about the bystander who has no good reason not to be impartial between the one and the five, when from an impartial perspective it is better if five survive than if one does? Suppose the trolley is headed for the five but the one is in front of them and his being hit will stop the trolley. Suppose he is unable to duck or run away, as he is permitted to do, but a bystander could easily move him off the track, though he could not move the five. This would save the one person but result in the five being hit. It would be impermissible for a bystander to move an inanimate object that blocked the path of the trolley to the five. Yet I think even an impartial bystander may and should help the one escape.¹⁹ One way of explaining why it can be permissible to save the one person and let the five be killed is that if there were nothing to be gained by the one person being hit, the bystander should certainly save him, and the fact that other people would benefit from the one person being hit (given that he has no duty to help them in this way) is not a sufficient reason for him to be left in that situation. It is true that helping the one would result in the five facing a threat from which they had previously been protected. But it would do so by removing the person who protects them when the reason to help him is not defeated by his usefulness in protecting them.

All this is consistent with its being right for an impartial bystander who cannot help everyone to remove the five off the track if he can, leaving the one to be hit by the trolley. In this case, he saves a greater number of people by means that do not require that the one be hit, though the one will foreseeably be hit because he is not saved. Removing the one person who would shield the five is also consistent with its being impermissible for an impartial bystander to remove the only person toward whom the trolley is headed when his being in a safe location will cause rocks to fall, killing five others. In this case, the bystander should refuse to offer easy assistance he should ordinarily give to the one person, because he avoids creating a threat to a greater number of people who were not being protected from the threat by the one person. Giving the one person a shield that would deflect the trolley toward five others whom he does not protect also should not be done because it will help cause a threat to a greater number of previously

unthreatened people, and not by removing protection provided to them by the one person.

Furthermore, an impartial bystander removing the one and so preventing him from stopping the trolley headed to the five when the five cannot be removed is at least consistent with turning the trolley away from the five, though it hits the one. Indeed, suppose the impartial bystander had moved the one person away from the trolley to the only safe location possible, expecting the five to be hit by the trolley. Then he newly discovers that he could redirect the trolley away from the five but only in the direction where the one had been moved for safety. It would be consistent for him to do this, for here the one's death would be the result of getting the trolley away from the five, not the result of omitting to help the one merely because he can be useful to protecting the five from the trolley.

The point is that the impartial nonconsequentialist bystander should favor keeping the greater number of people alive but only by means that do not treat other people improperly and create improper relations between the one and the five. Earlier, it was suggested that any PPH should be seen as having two parts, one that distinguishes different ways of harming innocents and a second part that applies these distinctions only when it is a question of greater good being produced at the cost of lesser harm. Therefore, it seems that if there is a correct PPH, in its entirety it would govern the acts of only impartial agents. (Note that a PPH does not directly speak to the question of not aiding, only harming. So it does not account for some of what I have said about an impartial agent rescuing the one person or instead rescuing the five people.) The impartial nonconsequentialist bystander may do a great deal more of what the five are permitted to do to help themselves than what the one is permitted to do to help himself. For example, the impartial bystander is permitted to move the five to safety and also to shield them even if their being on safe ground causes rocks to kill the one and the shield deflects the trolley to one. But, as I said, he should not shield the one (who does not provide protection against the trolley to the five) or move him to safe ground if doing so will kill five people. The question remains, as it did in our earlier discussions of the five and the one: if the impartial bystander is permitted to cause the death of the one person in these ways rather than let five die, why not by turning the trolley?

The table below provides a template for some issues covered in this discussion as they pertain to the trolley problem. I will not fill in the answers.

	5 themselves	1 himself	Partial Bystander	Impartial Bystander
Means				
Remove person(s)				
(a) stop protecting others				
(b) cause threat to others				
Shield person(s)				
Redirect threat				

NOTES

1. These are in addition to ones already considered in lecture 1.
2. All diagrams should be understood to allow a trolley, represented by a horizontal line, to be moved back from where it is and onto a different track.
3. See lecture 1's last section for some cases where who turns the trolley could matter and why.
4. This description of method follows the one in my *Intricate Ethics* (New York: Oxford University Press, 2007), 5.
5. For a more detailed but still incomplete discussion, see my *Intricate Ethics*, of which parts of section 3 in this lecture are a brief synopsis.
6. Judith Thomson, "Turning the Trolley," *Philosophy and Public Affairs* 36, no. 4 (2008): 359–74.
7. *Ibid.*, 373, 374.
8. I made this point in *Intricate Ethics* in response to comments by Derek Parfit.
9. For example, in "Harming Some to Save Others," *Philosophical Studies* 57, no. 3 (1989): 227–60; *Morality, Mortality*, vol. 2 (New York: Oxford University Press, 1996); and *Intricate Ethics*.
10. In *Intricate Ethics*, in discussing harm from mere means, I emphasize the difference between what we introduce into an environment having harmful properties and its triggering harm by entities independently in the environment.
11. I noted this in lecture 1 and will discuss it further below.
12. This formulation of the PPH speaks of whether in fact there are certain causal relations. But I believe the PPH should actually be formulated with a modal condition. That is, we should be concerned with whether, for example, a harm is required to produce an end or whether our act makes it the case that it is possible to produce an end without a harm if the harm (which actually produces the greater good) would not occur. I discuss this in *Intricate Ethics*.
13. See Judith Thomson, "The Trolley Problem," *Yale Law Journal* 94, no. 6 (1985): 1394–1415. In *Intricate Ethics* I also explained what I meant by ruling out mere means that cause harm "at least directly." I did not mean that a very complicated "Rube Goldberg" type of mere means causing harm would be permissible when greater good would result. Rather, I focused on whether what we introduce in an environment (complicated or not) to save some people itself causes harm by contrast to its causing harm by affecting what is in the environment independently

of what we introduce in trying to save these people. The former is ruled out by the PPH, while the latter need not be.

14. The one person's being hit is a necessary causal means to saving the five, but unlike standard cases in which we intend to bring about means to our end, in this case we take advantage of the causal role of a side effect (the one's being hit) of what we do (turn the trolley). It is said that a rational agent must intend the means to his end, but this need not be so when the means will come about as a side effect of what we have some other reason to do (turn the trolley away from one of its threatening positions). (Here is a wartime analogy: We bomb a munitions plant whose blowing up kills a few children as a foreseen side effect. We do not intend their deaths. However, there would be no point in our bombing mission if they did not die since it is only because they die that their parents are grieving and cannot quickly rebuild the factory. In this case, we do what destroys munitions only on condition of the causal role of a foreseen bad side effect, but it would be wrong to do anything besides what is necessary to bomb the plant in order to bring about the deaths of the children as a means to their parents' grief. I call this the Munitions Grief Case. It is neither a case of mere collateral (useless) harm nor a case of deliberately killing innocents to cause grief. For more on this, see chapter 4 of my *Intricate Ethics*.)
15. I discuss this in *Intricate Ethics*, 137, among other places.
16. I noted some of them in *Intricate Ethics* and draw upon that discussion in what follows.
17. Thomas Hurka emphasized this to me.
18. In previous work on this issue (in "Supererogation and Obligation," reprinted in *Morality, Mortality*, vol. 2), I described how supererogation may take precedence over a duty.
19. This is what I said in previous work about helping the one escape in the Loop Case where the one being hit would be useful in a somewhat different way. See, for example, the discussion in *Intricate Ethics*.