Responses to Commentators on *Intricate Ethics*[1]

F. M. KAMM
Harvard University

I

Some of the commentators on *Intricate Ethics* (IE) complain of my method. One finds the main ideas 'Kammouflaged' because the relevant causal distinctions are so fine-grained and the cases that illustrate them so numerous (Richardson, p. 82). Some say that they do not have the intuitions about many cases that I have, that I concoct dubious and ad hoc distinctions and invest them with moral significance; I am Ptolemaic in that new crystalline spheres and epicycles are constantly being added in an attempt to fix the appearances (Norcross, p. 74).

My only response to these complaints is that I start with the intuitions that I really have. The fact that my critics sometimes do not share my intuitions could account for why they would not attempt to unearth the distinctions I think are present. Rather than repress or prune my intuitions, as Richardson suggests, I choose to investigate them. I think I would compare this to an art critic who finds more value in one work of art than another, when many people see no difference between the two, and thinks he is on to something. The last thing he wants to do is make himself less sensitive. It is worth noting that the three critics in this symposium have a range of intuitive responses, from Alistair Norcross who does not have as fine-grained intuitions as Henry Richardson, to Michael Otsuka who has intuitive responses that are more fine-grained than Richardson. This will become clear below.

Once I have the intuitions, I seek to find what it is about the cases that is leading me to have the intuitions. I do not add 'epicycles in order to fix the intuitive appearance'; rather I try to find what it is in a case that I am responding to that leads to my intuition. (When the art critic finally isolates the brushwork in a certain area that is making him think the painting is so good, he is not 'adding' the brushwork.) I may discover that this is something about the way I am conceiving of the factors present in a case or the way I am focusing on the causal relations, or something else that accounts for my intuitive judgment. I do not necessarily intuitively find these underlying factors morally significant, though I may. (I certainly do not decide that action in a new

[1] F. M. Kamm, *Intricate Ethics* [IE] (New York, 2007).

© 2008 Cambridge University Press
doi:10.1017/S0953820807002944
Utilitas Vol. 20, No. 1, March 2008
Printed in the United Kingdom

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

case is (im)permissible because it contains a particular such factor, for it is always possible that the case is a counterexample to the importance of the factor. One should always check for one's intuitive judgment of a case because of this.) When I do not have an intuitive response to the found factor, I only claim that it accounts for my intuitions, and if my intuitions are morally correct, then these distinctions which generate them, perhaps surprisingly, have moral import. So, it may be that we do have to draw distinctions not drawn before in order to account for the appearances. Philosophers very often point to distinctions that make a difference though they were not noticed before.

Next, we might go deeper to investigate whether there are any ideas that we think have some intuitive moral significance with which these found distinctions are connected. I said in the introduction to my book that I would do much less of this sort of thing, focusing rather on considering my intuitions and what differences in the cases seem to generate them. I wrote barely two pages on the deeper issue, roughly making one suggestion about the difference between substitution and subordination, and at the end of that I even suggested that this might not be the right answer.

II

My commentators focus on chapters 4 and 5, but Norcross also discusses a part of chapter 1. Hence, I shall start with this part of his comments.

Norcross comments on my discussion of whether there is a moral difference per se between killing and letting die. His discussion is divided into three major questions. (I) Do the properties that I say distinguish between killing and letting die really do so only if we assume a doing/allowing distinction? (II) If certain properties that distinguish letting die from killing per se (by being essential to one but not the other) are morally *significant*, does this mean, as I claim, that the distinction between killing and letting die per se (rather than just between these properties) is morally significant? (III) Can properties behave differently in one context than in another? My response shall focus on (II) and (III).[2]

[2] Very briefly, I also note that Norcross thinks that the answer to question I is yes because he thinks that several properties that I claim are essential to letting die but not to killing could also be construed as essential to killing, unless we give weight to the doing/allowing distinction. The latter is a distinction that he finds metaphysically and morally problematic. First, he thinks that when an agent kills, as well as when he lets die, there is, at least potentially, a threat already present. This is because the killer could still present the threat. I had said, however, that the presence of (at least) a potential threat in the case of letting die requires an accomplice to the non-aider who would still bring the threat about initially but in Norcross's case the agent himself would have to bring about the threat. So the potential threat could exist, at least to begin with,

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

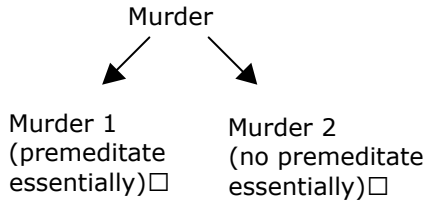


Figure 1.

Norcross thinks that the answer to (II) is no. He offers what he thinks are analogies to my argument for a 'yes' answer. Here is one analogy. It is an essential property of second-degree murder that there is no premeditation. It is an essential property of first-degree murder that it involves premeditation. It is not essential to murder (the class that includes both first- and second-degree murder) that it involves no premeditation. Not being premeditated is a morally relevant property that can make a murder with it morally better than a murder without it. However, that does not mean that second-degree murder (that has 'no premeditation' as an essential property) is morally better per se than murder which lacks 'no premeditation' as an essential property (though it may or may not have the property in particular instances). It is only better than first-degree murder which essentially has premeditation. Figure 1 above helps illustrate his argument (whose form repeats itself in his other analogy).

My concern about Norcross's counterargument analogy is that it does not accurately capture the form of my original argument. I said that if an essential property of letting die, when 'exported' in a case of killing, made that case of killing morally better than a case of killing lacking the

without the agent who would let die but not without the agent who would kill. Second, Norcross thinks that when an agent kills, as well as when he lets die, his victim only loses out on what he would have gotten from the agent's help; in the case of killing, this help is the agent refraining from killing. However, I would respond, in order to help his victim in this way, the killer must first present at least a potential threat from which he could refrain. But then this threat will be directed toward what his victim would have had independently of the agent's refraining, and this is what the victim will lose if he is killed. By contrast, the person who lets die need not present the initial threat to what someone had independently of him and so in not being aided, the victim loses out on only what he would have had with the agent's help. Finally, Norcross thinks that a potential killer is imposed on in being required to refrain from killing just as a person is imposed on in being required to refrain from letting die (i.e., in being required to give aid). But, I claimed, the killer is being imposed on defensively, to prevent his first imposing on someone else by killing. The person who lets die would be imposed on defensively only if letting someone die was also an imposition on the victim. Letting die, like killing, will determine that someone dies, but that does not mean that we impose on someone whenever our self-regarding conduct results in his life being shorter than it could otherwise be. (If it meant this, would we also impose on a potential killer, if we merely refrain from giving him what he needs in order to kill someone?)

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

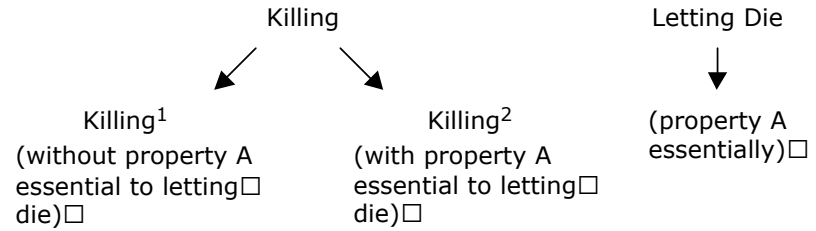


Figure 2.

property, then that was a reason to think that letting die was morally better per se than killing per se. (This was so given the additional conditions that (1) killing had no essential properties that could play a similar role in letting die cases, and (2) the essential property of letting die played the same role in its home ground of letting die). The figure above helps illustrate my account.

My argument did not claim, as Norcross's analogy would suggest, that killing[1] is morally better than killing (which I agree would make no sense). It claimed that some behavior, namely letting die (which is not merely a type of killing) could be shown to be better per se than killing, if it was one of letting die's essential properties that made killing[2] morally better than killing[1] that lacked property A. (This was so given the additional conditions (1) and (2) described above.)[3]

The comparable argument in Norcross's case would involve finding a non-murder behavior *x* that essentially has non-premeditation of something bad as one of its properties, and then claiming that it can be shown to be morally better than murder per se.

Here is another (fanciful) illustration of the argument form described in Figure 3:

If dogs had no essential properties that were better than cats' and affectionateness is a good property, then cats are better than dogs per se, though this does not mean that some dogs will not be as good as cats, in virtue of being more like what cats essentially are.

Now consider Norcross's discussion of III (i.e., can a property behave differently in one context than in another?). (a) He considers the two cases that James Rachels compares in order to show that killing and

[3] This is the part of my argument concerning killing and letting die that deals with what I called the 'exportable' essential properties of killing and letting die. I also said that essential properties of either killing or letting die that could never be present in (exported to) cases of the other type of behavior might also be used to show that there is a moral difference between killing and letting die per se.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

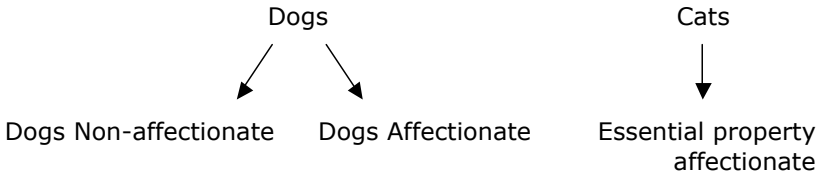


Figure 3.

letting die do not differ per se.[4] In one case, someone deliberately drowns a child and in the other someone deliberately lets a child drown when he could easily rescue him. Norcross agrees with Rachels that the killing and letting die are morally as bad in this context, where all contextual factors are held constant. (b) The further question is whether killing and letting die could behave differently (i.e., one makes a moral difference that the other does not make) in another context where all factors are held equal. The answer is important, because this could show that killing and letting die differ per se, even if this does not show up in Rachels's cases. (c) Norcross takes the view (which he attributes to me) (p. 69) 'that a factor such as the difference between killing and letting die could sometimes be morally significant and sometimes not' to be the claim 'that the factor's influence could vary in cases where *all other factors are equal*'. It is not, he thinks, the claim 'that other factors could sometimes cancel out the influence of the factor in question'. (d) He does not deny the possibility of such variance, but (p. 71) '[it] must be accompanied by an explanation of why the distinction is morally relevant in some cases and not in others'.

My response begins with noting that (in IE, chapter 1), I suggested that in Rachels's cases, killing and letting die still make a moral difference. I tried to bring this out by, for example, imagining what we could permissibly do to the person about to kill and the person about to let die in order to stop them from their respective wrong behaviors and save the child. I thought one could go so far as to kill the first person, if that were necessary to stop his killing. I did not think it would be permissible to kill the second person, even if the only way to save the child from drowning was for the dead body of the person who would have let him die to fall into the bathtub, thereby popping the child out. (I've relied on intuitive judgments just as Rachels did.) I agree that one should try to provide an explanation of difference (and equivalence). The explanation for the difference in the bad behaviors, I thought, was related to the fact that the killer would take away a life the child

[4] In Rachels' 'Active and Passive Euthanasia', *New England Journal of Medicine* 292.2 (1975), pp. 78–80.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

would have had independently of him, while the person who let die would not have done this. (In his discussion of his question I, Norcross argues that one cannot make sense of this explanation without relying on the doing/allowing distinction, and he doubts that that distinction has moral significance. I shall return to this point below. My point now is just to discuss my view of Rachels's cases and to suggest a possible explanation of a difference.)

Suppose, however, that killing and letting die make no moral difference in Rachels's cases. Could they still make a moral difference in other contexts? My first point is that I do not think, contrary to what Norcross says (in [c]), that there is a necessary contrast between (i) a factor's influence varying in cases where all other factors are equal and (ii) a factor's influence being cancelled out due to other factors. This is because if killing and letting die differ morally per se, then if we place them in equalized contexts, it is quite possible that some factor equally present in the context of both cases will cancel out the effect of the morally significant property in, for example killing, while having no other effect on the properties in letting die. An example of this is given by the context in which someone deserves capital punishment and we can bring about his death either by killing him painlessly or leaving him to die of a painless disease. Morally, the difference between killing and letting die may make no difference to which we do. The explanation of this in this case would be that we are entitled to take away the life of the criminal that he would have had independently of us. In another example, someone, A, starts his life growing in someone else's, B's, body, though B did nothing to bring this about. Suppose that A has no right to the use of B's body, even in order that he keep on living. Then B may be entitled to kill A to remove him, just as much as he would be entitled to let him die. The killing and the letting die in both these cases involve losing life A is not entitled to keep by way of using B's body, and this may explain why killing and letting die are not morally different in these cases. (These cases also make me wonder whether, as Norcross thinks, not letting die always involves *doing* something, as B would wind up providing assistance to A without doing anything.)

Norcross's view of Rachels's cases is that malevolent intention (as well as harm to the victim) in both the killing and letting die is the morally relevant property, and there is no moral difference whether malevolence is carried out by killing or letting die. He notes that some might still think that malevolence merely 'swamps' the moral difference between killing and letting die, but the difference will come out in a different equalized context. For example, when benevolence is the motive, as in euthanasia, would killing be morally wrong while letting die is permissible? Norcross thinks the morally relevant factor here is benevolence (and actual benefit to the person who dies), not whether

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

we bring about death by killing or letting die. However, Philippa Foot presented a case to show that this is not true.[5] It is in a wounded soldier's interest to die soon, otherwise he will certainly be tortured and killed by his captors. In one variation of the case, I, a fellow soldier (not a doctor), have medicine that can save his life so that he does not die before being captured. He wants the medicine. In the other variation, I can painlessly kill him, but he does not want to die. Foot believes I may permissibly refuse to give him my medicine, but it is impermissible to kill him against his wishes. I agree with Foot that this is another case that shows that there is a moral distinction between killing and letting die, for in this case when all other factors are equal, someone's refusal to die should prevent my killing him, but not necessarily my letting him die. This is presumably because he has a right not to be interfered with, even when it would be good for him, but my not aiding him does not interfere with him (though it determines his fate), and he has no independent right to what I would use to save his life.

III

Michael Otsuka discusses chapter 4. He agrees with me that there is a distinction between doing something in order that, or with the intention that, something else will occur and doing something because what we do will make something else occur. He thinks that I correctly apply this distinction in my Party Case. (This is a case in which I want to have a party but will not have it if I must clean up the mess afterwards. I then find that psychologists have shown that guests who have had a nice time will feel indebted and clean up after a party. I have the party only because I know that my guests will feel indebted and clean up, but I do not give the party in order to make them feel indebted in order that they clean up.) However, he does not think, contrary to what I believe, that the in-order-that/because-of distinction plays a role in the proper analysis of the Loop Case. (In the Loop Case, if a trolley is redirected from a track on which it will hit five people, it will only go onto another track that loops back to the five. It is only because there is a person on the looping track whose weight stops the trolley that the five are ultimately saved.)

(1)(a). He first considers (p. 97) one ground that I give for thinking that turning the trolley in the Loop Case need not indicate an intention to hit the one person. That is, suppose that doing only what is necessary to get the trolley away from the five as it is headed when we came on the scene fails to result in the one being hit in such a way that he stops the trolley. For example, the trolley could jump over him and on toward

[5] In her 'Euthanasia', *Philosophy & Public Affairs* 6.2 (Winter 1977), pp. 85–112.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

the five. It is not inconsistent of the turner to refuse to give an extra push just in order to get the trolley to hit and stick to the one. But a commitment to give such an extra push *would* follow from an intention to hit the one as a means to stopping the trolley, other things being equal.

Otsuka's strategy is to accept what I say about this Extra Push Case, namely that giving the extra push involves intending to hit the one, and show that what is done in the Extra Push Case is exactly the same as what is done in giving an extra push in another case I present, the Two Loop Case. In the Two Loop Case, one redirects the trolley from the five onto a looping track on which no one sits. One then gives an extra push to turn the trolley from this looping track headed to the five, though this just turns it to another track looping toward the five where, however, one person on the track stops it. But what is done in diverting from the first loop in the Two Loop Case is the same as what is done in the original Loop Case. Hence, he argues, if giving the extra push in the Extra Push Case involves intending to hit the one person, so does merely redirecting in the Loop Case, contrary to what I claimed.

As Otsuka analyzes the Extra Push Case, he says (p. 99),

> It does not follow . . . from the fact that such hitting of the one is causally necessary to prevent the trolley from jumping over him that hitting him is also causally necessary to get the trolley away from the first path that loops back towards the five (after the trolley jumps over him) . . . giving the trolley an extra push is in itself sufficient, independently of its hitting the one, to get the trolley onto a different path from the one it would otherwise have travelled.

That is, Otsuka believes that even if there is no other way to stop the trolley jumping over the one person besides giving it an extra push that would lead to its hitting and sticking to the one person, it is still true that changing the trajectory of the trolley by giving it the extra push is in itself sufficient to get the trolley onto a different path. This changing of the trajectory begins prior to the trolley's hitting the one (as it is by changing the trajectory that one does what makes the trolley hit the one). This is what makes it like the Two Loop Case, according to him.

(1)(b). My answer to this is that I had conceived of the Extra Push Case differently. I meant to contrast giving an extra push in the Two Loop Case and in the Extra Push Case. In the Two Loop Case, an extra push is sufficient to get the trolley onto a different path without hitting the one. Notice, crucially, that in giving the extra push in the Two Loop Case, nothing extra is done only because doing it is necessary to get the trolley subsequently to hit the one person on the second looping track. It is true that giving the extra push will result in the trolley hitting and sticking to the person on this second loop, but this does not mean that anything beyond what was necessary to turn the trolley away from its

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

first loop was done to produce the subsequent hit. It is a side effect. If some extra movement in the extra push were made only because it had this role of hitting the one, my view is that the agent would then be acting intending to hit the one person, rather than acting because the one person will be hit as a result of doing only what is necessary to turn the trolley from the first loop.

In the Extra Push Case, as I imagined it, the push was called extra because it was not needed to get the trolley away from its first hit on the five, but it was needed to get the trolley to hit and stick to the one person in a way that prevents its jumping over him. I had imagined that the extra push was given in the course of the original turning of the trolley from the five, but not just so as to redirect it from jumping over the one and onto a different trajectory where, as a side effect, it would hit and stick to the one (preventing the trolley from jumping over him and going on to hit five). The latter sort of push might make it like the extra push in the Two Loop Case, except that one acted in advance, with foresight to a second necessary redirection away from the five. If the extra push in the Extra Push Case merely redirects the trolley from looping as it does in the Two Loop Case, I would say that we could be redirecting because it will hit, but not necessarily in order to hit, the one.

According to Otsuka's conception, the extra push in the Extra Push Case is necessary and sufficient to get the trolley away from jumping over the person. That is, Otsuka agrees that hitting and sticking to the one is eventually necessary to stop the trolley, but he assumes this hitting and sticking will just happen as a side effect of a push that changes the trajectory of the trolley (over the person). But this should be an open question; it remains open whether anything extra must be done in this push that changes the trajectory specifically in order to get the trolley to hit and stick to the person so that it does not jump over him eventually anyway. It was *my point in constructing the Extra Push Case* to imagine a case in which this extra would be done for this additional purpose. (The extra push could involve an extra hard push, anything done especially to get it to hit and stick to the one.) I do not see why my conception of the Extra Push Case is not a possibility.[6]

Hence, it need not be true that the Two Loop Case and the Extra Push Case stand or fall together or that there are 'no good grounds to differentiate the extra pusher's intention in the Two Loop Case

[6] We could also imagine the Extra Push Case as not involving a trolley that might jump over the person. Rather, unless we gave an extra push (or made the push extra hard), the trolley would merely drag the person it hit onto the loop as it headed to the five. The extra push would change the impact on the one so that both person and trolley remain stationary.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

from his intention in the Extra Push Case' (p. 100). Only in Otsuka's conception of the Extra Push Case is this true. Furthermore, only under my conception of it would I maintain that it involves intending the hitting of the one.

2(a). Otsuka, at one time, also argued against my view that one may be acting only because one will hit a person and not in order to do so in the Loop Case by constructing the Five Loop Case.[7] It involves five tracks looping to the five people, but a person whose being hit can stop the trolley sits only on the fourth track. We have to make choices involving (1) sending the trolley to, and diverting it away from, several tracks that would loop back to the five and (2) sending the trolley to, but not diverting it away from, the fourth of five tracks on which one person will be hit, thereby stopping the trolley from hitting the five.[8] Apparently, Otsuka now accepts the following criticism of his original view and has omitted discussion of the Five Loop Case in his current comments. I include my discussion of his case because it is revealing of fundamental issues that are still in question.

As I understand it, Otsuka thought that it is artificial to say that you keep redirecting the trolley away from several looping tracks to prevent the hit on the five (just as you would redirect away in a standard trolley case). This is because you would not redirect away from the fourth track, on which the one person sits, to the fifth track, where no one sits. But my sense is that one would not redirect from the fourth to the fifth track because this will result in the five dying. So there is a reason not to redirect there. But there is a partial reason to redirect it from the third track to the fourth – to avoid its looping to the five – and it makes sense to act for that reason *because* the trolley would then go down the fourth track and be stopped by its hitting the one person.

Otsuka, by contrast, saw this type of analysis as 'an artificial construction of your reasons for action' (p. 106). He saw the Five Loop Case as involving our weaving our way to the fourth track in order to hit the one person as a means to stopping the trolley. Since the Five Loop Case is only a more complicated version of the simple Loop Case, Otsuka believed that it shows that in the latter, we also turn the trolley in order to hit the one person and not just because he will be hit.

2(b)(i). I would respond to these complaints as follows. First, I have suggested in IE that we can test whether someone turns the trolley intending to hit the one by seeing (1) if he would give an extra push (as I have described above my conception of it in the Extra Push Case),

[7] In his original comments delivered at the APA, Pacific Division, March 2007. My page numbers throughout Section 2(a) *only* refer to that original document.

[8] Insofar as it involves choices about to which track to send the trolley, it is similar to but more complicated than my Double Track Case discussed in my *Morality, Mortality*, vol. 2 (New York, 1996) and in IE.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

and (2) if he would fail to rescue the one person on the side track if this were easily done. (Otsuka does not discuss the Rescue Test.) These tests can be applied to the Loop Case and the Five Loop Case. Consider the Rescue Test.[9] I think it is consistent with no change of intention in a rational agent for him to turn the trolley away from the five in the Loop Case, and also up to but not further than the fourth track in the Five Loop Case, because the one will be hit, and yet to help the one person off the track if this suddenly becomes possible, even though this means that the five will then be killed. If he refrained from rescuing when this cost him nothing, this would indicate that he intended the death of the one to save the five. By contrast, it would be inconsistent for someone who had turned the trolley in order to hit the one and retained that intention to then try to rescue him. Similarly, someone who threw a person off the bridge in order that the person fall in front of the trolley and stop it (Bridge Case) would show a change of intention if he then rescued the one.

Consider the Rescue Test in an analogous not-aiding (versus harming) case. In a version of the so-called Transplant Case involving letting die, we intend the death of someone as a means to getting his organs so that we can save five people with them. Hence, we do not give him a life-saving medicine to which he has a right. In what I call the Secondary Transplant Case, we direct a scarce medicine to save five from disease A, thereby leaving one other person to die from disease A. We do this even though we know that the medicine will create a side effect that would be fatal to the five were it not that the organs of the one who is left to die without the medicine will be available after his death for transplantation to the five. In the Secondary Transplant Case, we save the five with the medicine only because the one will die; it would be a waste of time giving them the medicine, as they would die of the side effect anyway if the one's organs were not available. (This is analogous to its being a waste of time to turn the trolley from hitting the five if the one were not there to be hit and stop the trolley from looping.) But suppose that a new cure that could save only the one (not the five) from disease A became available after we had given the other medicine to the five. It is consistent with our having given the medicine to the five only because (we thought) the one would die, to now try to save the one even though this means the five will die. We would do this if we wish to avoid denying him a new medicine that the five cannot use only in order that he die.

2(b)(ii). An additional response to the complaint embodied in the Five Loop Case is to consider a case where Otsuka agrees that we

[9] Note that, even now, when Otsuka no longer argues using the Five Loop Case, he agrees (in correspondence) that 'the Rescue Test can still be deployed against my claim that the agent intends the hitting of the one in the Loop Case'.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

can be acting because we will hit someone without acting in order to hit him and see if this case would also fall prey to the supposed counterargument provided by the Five Loop Case. Otsuka believes that in his Six Behind One Case, it can make sense to say that we turn the trolley away from five people onto a track where one person sits in front of six others *because* the one person will be hit, thereby stopping the trolley from going on to kill the six, but not in order to hit him. This is so even though we would lose our justification for acting if the one were not hit because then six would be killed instead of five. Indeed, he uses the same sort of language to describe this case that I use to describe the Loop Case and other cases in chapter 4 of IE: We do not do an act in order to create an effect (the one being hit) that will defeat a potential bad side effect of our act (the six being hit) that would rob us of a justification for action. We must have an independent primary (goal) reason for action (saving the five) and that is what we act in order to bring about.[10] (Also, compare an additional point about the Loop Case based on what I say in chapter 5. Abstract not only from the hit to the one but also from all further effects of what we do to remove the trolley threat as it is when we come on the scene. Then we can see that in turning the trolley, we do something in order to save the five from the threat as it was when we came on the scene because the looping will be stopped by the hitting of the one, just as the hitting of the six will be stopped by the hitting of the one in Otsuka's case.)

Now, consider my variant on the Six Behind One Case that is like the Five Loop Case. There are five tracks going from the five people toward the six people, each track allowing a diversion to the next one. Only on the fourth track is the one person who will stop the trolley if he is hit. Consider someone who thinks that turning up to the fourth track in the Five Loop Case indicates the intention to hit the one. Would he say that if we keep on diverting the trolley away from the six no further than the fourth track, we are also turning with the intention to hit the one person in this case? But then will he also have to conclude that the original Six Behind One Case involves intending to hit the one rather than merely turning away from the five because the one would be hit? If he will not conclude this about the Six Behind One Case on the basis of the variant that is like the Five Loop Case, he should not conclude it about turning in the original Loop Case on the basis of his view of the Five Loop Case.[11]

[10] In discussing my Store Cases I and II in IE, I try to distinguish between refusing to act unless we achieve two goals and refusing to act unless a potential bad effect of our achieving our one goal will be defeated. I shall return to the second of these cases below.

[11] Otsuka thinks that in both the Five Loop Case and the analogous variant of his Six Behind One Case, 'the agent does intend the hitting of the one'. (I disagree.) But he does not think that the agent intends the hitting of the one in his original Six Behind One Case. This is what led him to conclude that he could not infer that an agent in the

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

We might also consider another type of condition on action, namely a side constraint, that may further clarify the Five Loop Case. (Note that Otsuka says (p. 102, n. 17) that not killing the six might play this role of a side constraint on saving the five in his Six Behind One Case.) The case I discuss in IE is one where I intend to pick up a glass in order to drink some water, but will do so only if doing this does not harm you. However, I do not do the act in order to do something that does not harm you; that is not a goal of my action. Suppose there are different ways in which I could raise my glass, all of which will harm you except one, and so I choose to do that one. It is still not true that I perform my act in order not to harm you, even though I select among the acts on the basis of which will not harm you. Indeed, if there were a moral prohibition on doing an act in order to create an example of not harming someone, we would not have violated this prohibition if we select to drink only in the way that does not harm someone.

2(c). Our discussion of the Five Loop Case led us to the Six Behind One Case. But independently of its relation to the Five Loop Case, Otsuka uses it to present a case where he thinks that we can correctly say that we act because we will hit the one person but not in order to hit him. He contrasts it with the Loop Case, where he thinks it is inappropriate to use the same language. In the Six Behind One Case, he says, our goal (to save the five) has been completely accomplished; the hit of the one prevents the six from being killed. As noted above, he does not think it was our goal in acting to save the six, as one does not do something in order to mitigate the bad consequences of doing that very thing. By contrast, in the Loop Case, he says (p. 102), our goal is to save the five, and the hitting of the one serves the goal of (is a means to) saving the five, hence one must be acting in order to get the trolley to hit the five.

I considered such an argument in chapters 4 and 5 of IE. In chapter 5, I argued that it made a moral difference whether the hitting of the one served our goal of saving the five (1) by *sustaining* what we had done without it or (2) by *producing* an additional component of our goal. As I saw the Loop Case, abstracting from further effects of our turning the trolley, we have achieved the goal of saving the five when we turn the trolley. But, of course, what we have achieved could be undermined by the further effects of what we have done, so abstracting from these further effects, we have achieved only (what I called) the 'structural equivalent of our goal'. Still, this contrasts, I thought, with the Tractor Case, where both a trolley and a tractor threaten the five,

original Loop Case intends to hit the one on the basis of his views about the Five Loop Case. Hence, he dropped the Five Loop Case as an argument against my view that the agent in the Loop Case need not intend to hit the one. (My remarks here are based on Otsuka's correspondence.)

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

and the tractor faces them independently of any effects of what we do to redirect the trolley. Abstracting from further effects of turning the trolley leaves us with five people still threatened by a tractor. I thought of this case as one in which we would need to produce another component of our goal of saving the five, namely get the tractor away. By contrast, in the Loop Case, additional effects of what we have done both introduce a new threat (looping of the trolley) and defeat it (by hitting the one), thus sustaining the structural equivalent of our goal. (To the extent to which I distinguish between sustaining (in Loop) and producing (in Tractor) a goal, I, like Otsuka, thought it was important whether we *produce* our goal independently of means involving harm to the one person.)

In chapter 4, I also argued that it is not correct to assume that a rational agent must always intend what he knows to be a necessary means to his goal. Hence, contrary to what Otsuka says, we cannot show that someone intends an effect of his act and acts in order to bring it about just by showing that he knows that it is a necessary means to a goal that he intends. For example, suppose that our goal is to build a house. We know that we cannot do it unless we have both bricks and dynamite. Suppose it is wrong to buy dynamite for this project. However, it turns out that a side effect of my buying bricks is that the store gives out a bonus supply of dynamite. I need not intend to acquire the dynamite, given that it will come as a side effect. So I proceed with my goal of building the house.[12]

In the Store Case II, presented in chapter 4, I described someone who intends to get a pair of shoes that he can keep to wear. He is poor, however, and if he drives to the store and back, the gas bill will be so great that he will have to sell the shoes to pay the bill, thereby defeating his goal of getting and keeping the shoes. He learns that the store has a special offer: they will repay transportation costs of anybody who buys a new pair of shoes. The person is very proud and would never do anything especially in order to get the store to repay his transportation

[12] I also noted in chapter 4 of IE, that it might be possible (unrealistically) to construe the intentions of the person who redirected the trolley in the Loop Case as follows. He intends merely to get the trolley away from hitting as it was headed independently of his actions. He does this, however, only *because* he foresees that the one will be hit by the looping trolley and so save the five. That is, he sees the eventual death of the five as a bad side effect of his completely achieving his very limited goal (of getting the trolley away from hitting as it was headed independently of his actions) and that bad side effect would rule out pursuing his goal, because the death of one person is not justified by the achievement of his limited goal alone. However, the one's death can be compensated for by the good effect of the five eventually being saved. I made this point when considering whether the greater good that we cause must always be intended if it is to be able to justify our causing a lesser evil as a side effect. My answer was 'no', because a merely foreseen greater good side effect could also outweigh a lesser evil.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

costs, but given that their doing so is a mere side effect of his getting his shoes, he goes to the store. In this case, too, having the gas bill paid is a necessary means to his goal of keeping the shoes, but he does not act in order to get his gas bill paid. I said both (1) that it would be odd to say that he went to the store in part in order to get something that compensates for a bad side effect (gas bill) that only arises because he goes to the store and (2) that he need not act intending the repayment, though it is a means to his goal, when it is a foreseen side effect of something else he does.

2(d). In concluding my examination of the Six Behind One Case, I wish to note in passing that the fact that Otsuka believes that it is permissible to turn the trolley in the Six Behind One Case shows that he disagrees with Richardson. For Richardson would argue (as we shall see below) that we may not act when harm that we would cause to another is necessary in order for us to have sufficient justification for action, as in the Six Behind One Case. Here we see that Otsuka draws a finer distinction than Richardson, an instance that confirms my view that my commentators are on a spectrum when it comes to sensitivity (for better or worse) in intuitive judgments.

(Richardson also disputes (p. 89) whether there is a difference between harm being a necessary cause and an actual cause to an end, so he says (in describing his view) that '[t]he core idea of this account of inviolability . . . hinges on whether or not someone's being harmed is a causal means to the greater good'. This implies that he should think it is impermissible to bomb a munitions factory in wartime when destroying munitions *could* end the war, if the populace *actually is caused* to surrender by its terror over the deaths of some of its children, a proportional side effect of the bombing. Judith Thomson, who describes such a case, argues[13] that this would be the wrong conclusion, and I agree. My view is that this is the wrong conclusion because the harm's causing the surrender is not *necessary* to justify (or even explain) our acting, given that the harm (that does produce the surrender) is not necessary in order to produce the surrender. This is why I say that the use of modal notions should be involved in a principle describing when causing harm is and is not permissible.[14])

3. Otsuka's next set of criticisms are all related to his view that turning the trolley in the Loop Case cannot be distinguished morally from what we do in the original Bridge Case, in which we move a device

[13] In her 'Self-Defense', *Philosophy & Public Affairs* 94.6 (1991), pp. 283–310.

[14] See IE, chapter 5, for more on this. Norcross similarly misses my emphasis on modality when he describes my view. He incorrectly says (p. 72): 'It turns out that the exact details of the causal processes leading to the greater good and lesser evil are crucial for determining when harm is permissible and when it isn't.' It is important to distinguish actual causal processes from necessary (required) ones.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

in order to push someone over a bridge so that his falling in front of the trolley will stop it.

3(a). He first presents what he calls the Loop-Bridge Case, in which turning the trolley away from the five would only make it turn in a circle back to where it started, were it not that in its circular course the trolley triggers a lever that makes a pole push one person off a bridge and into the path of the trolley.[15] He is correct to say that I believe turning the trolley in this case is permissible, even though it involves pushing the person into the trolley. (The Loop Case involves pushing a trolley into the person. I had previously argued in my Lazy Susan Case that it is permissible to turn five people away from a trolley, even if this results in the turning Lazy Susan pushing one other person into the trolley.) He thinks, however, that it is difficult to see how the fact that the pole gets pushed as a result of the trolley being turned away from the one on the circular path could make pushing the man over permissible when it is impermissible in the Bridge Case.

Furthermore, he thinks that in the Loop-Bridge Case, one would redirect the trolley intending to hit the one. He thinks that he can demonstrate the latter point by comparing the Loop-Bridge Case with his Illuminated Trolley Case (p. 105), in which 'your pushing the lever causes the trolley to be illuminated by a spotlight as it whizzes down the main track, and this light's reflection off the shiny surface of the trolley activates a sensor that causes the person to fall from the bridge'. He claims that someone might mistakenly describe this case as follows. We eliminate the first threat of an *un*illuminated trolley headed to the five by illuminating it, knowing that this light source will trigger a sensor that causes someone to topple from a bridge in front of the trolley, thereby stopping the second problem we have created of an illuminated trolley headed to the five. He thinks that this description is analogous to what I say about the Loop Case: We eliminate the first threat of a trolley headed to the five by our diverting the trolley, knowing that this will result in hitting the one, thereby stopping the second problem we have created of a looping trolley. He says (p. 106):

> It is not as if you have reason, independent of the fact that the illumination will cause the one to fall in its path, to make it the case that an unilluminated trolley does not hit the five by illuminating it. Similarly, in the Loop-Bridge Case you have no reason to divert the trolley down the circular looping track except in order to eliminate the threat the trolley poses to the five by getting it to come to a halt by hitting the one before it hits the five. In abstraction from the fact that it would come to a halt by hitting the one as a result of its

[15] This case builds on (a) my case (IE, p. 124, n. 9), where turning the trolley away from the five would just send it in a circle back to where it started, were it not that letting one person on the circular track stops the trolley and (b) Derek Parfit's suggestion that turning the trolley causes the person to fall from the bridge.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

diversion, you do not eliminate or lessen the threat that the trolley poses by sending it on a circular detour.

(Norcross presents a case similar to Otsuka's Illuminated Trolley Case, except that it involves changing the speeds of a trolley – stopping the first threat of a slow trolley by speeding it up because we know this will lead to the speeding trolley hitting one person, which stops the trolley from hitting the five. Richardson also suggests a case (p. 89) in which we paint the trolley green and say that we have saved someone from being hit by a red trolley. So, my response to Otsuka is a response to Norcross and Richardson, too.)

3(b). I begin my response by saying that I do not think that the Illuminated Trolley and the Loop Case are analogous. The absence of illumination, which characteristic we try to eliminate in the Illuminated Trolley Case, is not a threat to the five. Looked at another way, changing the lighting (or slow speed or red color) of the trolley would not in itself have a point, independent of further effects, if we want to save the five. But a trolley hitting them from its initial direction in the Loop-Bridge Case is a threat and turning it away has a point if we want to save them, even if the point could be defeated by the further effects of turning the trolley. Otsuka says that we do not eliminate the threat the trolley poses by sending it in a circular (or looping) track. But that does not mean that in removing it from its initial hit, we do not deal with an aspect of its threat that, independent of anything else that happens, must be dealt with if the five are to survive. (Saying that there is a real point in diverting the trolley from its initial hit is not the same as describing one's intention in diverting as only ever to get the trolley away from hitting the five via the first path that it would otherwise have traveled. One can also have the intention to save the five completely. However, the point is that one need not always intend all the effects of one's acts that are necessary to bring about this final end of the five being saved completely. This is why, as described above, I also argue in chapter 4 that, in general, a rational agent need not intend the means that he knows are necessary to his end, unless intending them is necessary to producing the means. When these means occur as side effects, one need not intend them, even though one intends the end to which they lead.)

So, in the Loop-Bridge Case, by contrast with the Bridge Case, the one is toppled as a result of removing a threat to the five (in the sense just described). Contrary to what Otsuka says, I think that this difference in the causal pathway that leads to the toppling of the person from the bridge in the Loop-Bridge Case, by contrast to the Bridge Case or the Illuminated Trolley Case, is morally important. A closer (though not perfect) analogy to the circling trolley would exist if the trolley headed

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

to the five also gave out death rays to the five and redirecting these rays, so that the five will not die of that threat at least, caused someone to fall from the bridge and stop the trolley. (Call this the Death Ray Trolley Case.) In this case, as in the Loop-Bridge Case (in my view), there are two threats (actually or potentially), and it only pays to take care of one of them because doing so has the effect of stopping the other one.[16]

Otsuka says that just as illuminating the trolley is a mere means to pushing someone off the bridge, redirecting the trolley in the Loop-Bridge Case is a mere means to hitting someone. He concludes that we illuminate the trolley and turn it in the Loop-Bridge Case in order to hit the one person, not just *because* the one person will be hit. But again consider the Death Ray Trolley Case. It is true that independent of its causing the one to drop from the bridge, any point in stopping the death rays is defeated if the trolley would hit and kill the five anyway. But that does not mean that stopping something that must be stopped from causing harm if the five are to have any chance at life – either death rays or a trolley coming in one direction – is a mere means to stopping the trolley by way of an intention to hit the one person. Unlike pushing someone over a bridge in the Bridge Case, turning the trolley in the Loop-Bridge Case (or the rays in the Death Ray Trolley Case) has as its flip side the five being free of a form of a threat that had to be dealt with, whatever else had to be done or to happen, if they were to survive.

Furthermore, if diversion of the trolley were a mere contrivance and it was 'valuable merely as a means to make [the trolley] come to a halt' (as Otsuka says on p. 100), those who divert it should be willing to divert a second trolley that is not threatening anyone, so that it hits one person sitting on a side track, if his being hit will (miraculously) stop the first trolley headed around a loop toward the five (Two Trolley Case). But in this Two Trolley Case, though the second trolley would be headed to the one person, it is crucial that it would not also be headed *away* from a hit on the five, unlike the trolley in the Loop-Bridge Case. This contributes to the impermissibility of using the second trolley to cause the hit, just as the fact that pushing someone off the bridge in the Bridge Case is not the result of a threat moving away from the five

[16] Similarly, in Norcross's case (p. 75) in which we speed up a trolley, speeding up the trolley, independent of its effect of hitting the one person, has no relevance to saving the five. However, we could imagine a revised case in which only if we speed up the trolley would it jump over the five. Its speeding would also make it hit a person after it jumps whom it would not otherwise have hit. His being hit, while not necessary to stop the trolley from hitting the five in the front, does mean that it cannot loop back to hit them from the rear, as it otherwise would have. In this Speeding Trolley Case Revised, there is an independent point to speeding up the trolley – getting it to jump over the five – even if this point would be defeated by the further consequence of looping were the one not hit.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

contributes to its impermissibility. Note also, if we did not additionally turn the first trolley away from the five in the Two Trolley Case, the fact that hitting the one person by the second trolley was useful in stopping the first trolley from looping would not be sufficient to save the five, for they would die from the first trolley hitting them without its looping. This helps show that turning the first trolley away in the Two Trolley Case, and turning the only trolley in the Loop-Bridge Case, is more than a means to hitting the one. Because the Loop-Bridge Case is in this way different from the Illuminated Trolley Case, I also do not think the latter case shows that we redirect the trolley intending to topple the one from the bridge, rather than because it will topple him, in the former case.

My response to Richardson's Bypass Case is similar. In the Bypass Case, the trolley headed to the five can be directed to a bypass where one person's being hit would stop it from going back to the original track from whence it would have hit the five from the front. He says that in the Bypass Case, contrary to what I say in justifying redirecting in the Loop Case, 'we can no longer say . . . abstracting from the further consequences that include the evil*, that we have saved the five from being hit from the front' (p. 88).[17] This case is like the one that I discussed in IE,[18] where the trolley headed to the five can only be redirected in a circle that would end with the trolley being back at the point where it started (headed to the front of the five), were it not that hitting one person on the circle track will stop it. Intuitively, I believe that redirecting in this case of mine, the Loop Case, and the Bypass Case are all permissible. About all of them we can say that, abstracting from the further consequences that include the hit, there is a point in turning the trolley for it saves the five from a problem that faced them, a problem from which we must save them regardless of what else we must do or what else must occur, if they are to be saved. And this is true even if we would create another instance of this very same problem, if hitting the one person did not to occur and prevent it. By contrast, I think it would be impermissible to turn a second trolley threatening no one so that it hits a bystander when this hit would (somehow) stop the first trolley from hitting the five. Richardson's view, like Otsuka's, implies that there is no moral difference between the latter case and the Loop, Loop-Bridge or Bypass Cases. I disagree.

Notice that we can also create a Bridge-Loop Six Behind One Case: The trolley is turned from the five and it causes a pole to push someone off a bridge, and only his falling in front of the trolley stops six others from being hit. Presumably, Otsuka would find it permissible to turn the

[17] 'Evil*' is my phrase for 'evil or involvement leading to evil'.
[18] p. 124, n. 9 (also described herein in n. 14).

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

trolley in this case, even though without our pushing the one, we would not be justified in turning the trolley. By contrast, pushing someone off the bridge in the ordinary way (as in the Bridge Case) to stop the trolley headed to the six would be judged impermissible. Richardson, by contrast, should find both these cases impermissible as they both involve the one person's being hit causing an effect (the six saved) that is needed to justify our act.

3(c). To his set of cases, Otsuka next adds what he calls the Ramp Case, in which

> you can stop the trolley from hitting the five by opening a drawbridge. So doing will send the trolley up the ramp of the open drawbridge and into the air, where it will hit one person standing on a pedestrian bridge overhead. The trolley's hitting the one will stop its forward progression and cause it to fall into the water below the drawbridge. The trolley's hitting the one will also kill him. If, however, the one had not been overhead, then the trolley would have managed to fly over the gap, land on the downward sloping ramp, and hit and kill the five.[19]

He says,

> Given our intuition that it is impermissible to kill the one in the Bridge Case, we're therefore inclined towards the view that it is impermissible in the Ramp Case to send the trolley up the ramp and into the person on the pedestrian bridge. . . . If it is impermissible to launch the trolley upwards towards the one in order to stop it in this Ramp Case, then how could it be permissible to divert the trolley towards the one in order to stop it in the Loop Case? (pp. 107–8)

I agree that the Ramp Case is morally like the Loop Case, if we must move the trolley away from the five to set it to go up the ramp, but I do not think it is morally like the Bridge Case. Intuitively, I think that the fact that the one is pushed off the bridge as a consequence of moving the trolley away from its initial hit on the five makes a significant moral difference from what is done in the Bridge Case and makes turning in the Ramp Case permissible. It also means that we need not be diverting the trolley 'in order to stop it' in the Ramp Case, rather than just *because* it will be stopped.

I conclude that Otsuka does not show that the Loop Case is morally like the Bridge Case.

4. In concluding, Otsuka reflects on how intuitive responses to cases come about. He believes that if, historically, we first think about redirecting in the Trolley Case and judge it to be permissible, it will be hard to see how an additional piece of track could make turning the trolley in the Loop Case impermissible. However, if, historically, we first think about the Bridge Case and judge pushing the person to be impermissible, we will have a hard time seeing what differences

[19] He attributes the case to J. M. Fischer.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

make action in the Loop-Bridge and the Loop Cases permissible. (He reports student surveys showing this.) Alastair Norcross makes a similar methodological claim (p. 67) when he says (based on student surveys) that if, historically, we first think about killing someone for his organs to save five others in the Transplant Case, we will be less inclined to think turning the Trolley is permissible, and if we start with the Trolley Case, we will be more inclined to think that proceeding in the Transplant Case is permissible.

Given that, in Otsuka's view, relying on intuitions can lead us to be pulled to identify the Loop Case with the Trolley Case, but also with the Bridge Case, he thinks that we must decide how to think of the anomalous Loop Case on the basis of 'the cost' of giving up the idea that there is a 'morally significant distinction between treating as a means and not so treating' (p. 109). He thinks we have support for this distinction in a wealth of intuitive responses in cases and in 'strong theoretical grounds' (p. 110).

In response to Norcross's point, I would note that at least in the philosophical literature, it was the Transplant Case that came first and the Trolley Case second.[20] Contrary to Norcross's prediction, despite prior exposure to the Transplant Case in which they thought that killing was impermissible, many philosophers found that they thought that turning the trolley in the Trolley Case was permissible. Perhaps these philosophers were able to focus on each case separately, not allowing a desire for false consistency to impair what I would call their 'intuitive integrity'. But suppose Norcross were right. Then given that Otsuka is tempted by the relevance of the order in which cases are presented, would he be willing to conclude that the permissibility of turning the trolley in the Trolley Case is also an illusion? Perhaps not, because he, unlike Norcross, would argue that Transplant and Trolley Cases are distinguishable by the moral relevance of whether we treat a person as a (mere) means (in Transplant), for which he thinks there is much independent support.

My response to Otsuka is that I do not think that my attempt to distinguish between (a) the Bridge Case and (b) the Loop-Bridge, Ramp, and Loop Cases is unsuccessful (either as presented in chapter 4 or 5 of IE). Hence, I believe that even if one had started with a case ordering going from (a) to (b), if one had maintained one's 'intuitive integrity', one would have responded to the factors that distinguish them with different intuitive moral judgments. I also do not think that the reason one thinks that turning in the Loop Case is permissible, if one starts with the Trolley Case, is that one cannot see how an extra piece of track

[20] Philippa Foot's 'The Problem of Abortion and the Doctrine of Double Effect', *Oxford Review* 5 (1967), pp. 5–15; reprinted in *Killing and Letting Die*, ed. Bonnie Steinbock and Alastair Norcross, 2nd edn. (New York, 1994), pp. 266–79.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

could make a difference. This is an overintellectualized account of the grounds for thinking it is permissible to turn in the Loop Case. I have no a priori idea of whether an additional piece of track could or could not alter moral permissibility (by altering relations between people). Rather, I have an intuitive response to the Loop Case, not derived from any prior view about the possible significance of the piece of track.

Finally, what about 'the cost' of giving up the view that whether one treats someone as a means is morally crucial? (The common view is that it is wrong to use someone as a *mere* means, not that it is wrong to use them as a means. I shall henceforth assume that Otsuka is concerned with that common view.) First, notice that Otsuka himself, in his Six Behind One Case, accepts the permissibility of acting in another 'anomalous' case, in which it seems that we must use someone as a mere means to cause an outcome that is necessary to justify our pursuing a goal. He is open to this, while Richardson is not. Second, Otsuka cites Warren Quinn's view as the theoretical grounds supporting a prohibition on using as a mere means. In chapter 3 of IE, I critically examine Quinn's views on intending harm and using as a mere means. I will here only mention Quinn's concern that we not treat people as *available* for our use. I think that, at least, when the use of a person comes about as an effect of our doing something that either fully accomplishes a goal (as in the Six Behind One Case) or accomplishes an outcome that need only be sustained (as in the Loop Case), we do not treat the person as *available*, as there are very restricted ways in which he can come to be put into the position of being a means. We do not do something, in particular, whose only role is to put him in that position. Further, many cases that give intuitive support to the common objection to using as a mere means do involve putting someone in the position of being used as a mere means by doing something whose only role it is to do that (e.g. pushing someone off a bridge, as in the Bridge Case).

Unlike Otsuka, I think we should be open to the possibility that generalizations, such as that we must not use people as mere means, have been based on an incomplete consideration of cases and that a correct moral principle will take a somewhat different, though possibly related, form. (Being open in this way can be done without revising any intuitive judgments about cases that have been used to support the other generalization.) I also do not believe that we find correct principles by noticing that our intuitions pull in different directions and concluding that we must decide on the basis of 'cost' which way to go. This is to give up on moral *discovery* and substitute for it moral construction in which we, supposedly, have optional choices. Quinn himself seemed to reject such a method.[21]

[21] See his 'Action, Intention, and Consequences: The Doctrine of Doing and Allowing', *Philosophy & Public Affairs* 18.4 (Autumn 1989), pp. 334–51.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

IV

1. Richardson and Norcross discuss (mostly) chapter 5. Let me describe the origins of the view I am proposing there. Philippa Foot noted[22] that if we had to perform life-saving surgery on five people with a gas that would kill a bystander, we should not do it (Gas Case). If driving down a road to a hospital with five people to be saved would involve us in running over someone else, we should not do it. This is true even though the bad side effects were not intended and we were not using the bystanders as mere means to our end. (How will someone like Richardson explain the impermissibility of these acts, if he just focuses on not treating people as mere means? Having a simple theory such as his is not good if it is not explanatory.) But Foot also said that this did not mean that we should never aid five people at the cost of causing someone's death, for we may do so in the Trolley Case. (However, in Foot's version of the Trolley Case, it is the conductor who must turn the trolley. It was Judith Thomson who first argued[23] that a bystander may do so as well. This is now the standard Trolley Case.) Why is it permissible for us to turn the trolley and cause a death but not use the gas? To answer this question, I thought of revising the Gas Case in the following way. Suppose we saved the five people by surgery that harmed no one but their being alive and breathing normally changed air currents, and so fatal germs in the atmosphere that would otherwise not have gone to a bystander did. It would still be permissible to do the surgery, even if we knew what would happen. I suggested that this was because the greater good of the five being saved was causing the lesser evil, not the gas which was a mere means that caused the five to be saved.

However, it is also true that in this Revised Gas Case, the greater good more indirectly causes the lesser evil than the gas would in the original Gas Case. Maybe that was important because it signaled that it was not something that we *introduce* into the scene (like a gas) that harmed someone but rather the interaction between what we did (saving the five) and something in the environment independently of what we do (germs). (Of course, indirectness of causing a bad effect would not count morally if we *introduce* the indirectness, for example, by creating a Rube Goldberg device.[24] It would also not count if bringing about greater good were not at stake.) However, I then thought about the Trolley Case, and it seemed to me that if I could not move the trolley away from the five but I could move the five away from the trolley (by

[22] In her 'The Doctrine of Double Effect and the Problem of Abortion'.

[23] In her 'Killing, Letting Die, and the Trolley Problem', *The Monist* 59.2 (April 1976), pp. 204–17.

[24] So, contrary to what Norcross says (p. 79), I am not concerned, per se, 'with how many links there are in the causal chains connecting such harms and benefits to each other or to an agent's actions'.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

turning some large swivel table – a Lazy Susan device – that they were on), it would be permissible to do this, even if in turning, they crashed into and killed one bystander. Here the greater good of their being out of harm's way would directly, not indirectly, cause the lesser evil, yet it would be permissible to act.

Of course, in the original trolley case, it is a redirected trolley, not the greater good of the five being out of harm's way, that would directly cause the lesser evil. Why was this not like the gas causing the lesser evil in Foot's Gas Case? I focused on the fact that the greater good of five being out of harm's way was just the *flip side* of the trolley moving away; it was not that the trolley moving away caused something else, like their being pushed away from the trolley. It seemed to me that this made the redirected trolley causing the lesser evil almost like the greater good itself causing the lesser evil and that this was part of why we thought redirecting was permissible. Contrary to what Norcross says, this is not a matter of the certainty of the good coming about. It is a matter of whether something else bad has to happen, in addition to the trolley being away, or nothing else bad has to happen, in order for the five to be okay. So this is how my theory began before I started dealing with all sorts of other variations (such as the Loop Case) and trying to understand why I had certain intuitive responses to them.

It is true that I use the distinction between a non-causal flip side and ordinary causation to draw one of the distinctions in my view, and I talk of direct and indirect causation. I do not believe, however, that my distinctions are wedded to any particular metaphysical view, as Norcross claims (pp. 69, 72). For example, I would be happy to think of the first distinction as a distinction within causation. I am only concerned that the difference show up in some description of events.

2. Now to specific criticisms raised by Norcross and Richardson. Richardson claims that another distinction to which I point – the producing/sustaining distinction – does not have moral significance. He compares two of my cases. In one, we would turn the trolley away from five, who also have fatal organ failure, toward another track where a bystander will be killed by the trolley. His organs can then be transplanted into the five and save them. In the second, our turning the trolley from the five is what (somehow) causes their organ failure, and the organs of the one who will be hit can be used to save them. I suggested that there may be moral problems with turning the trolley only in the first case. This is because in the second case, turning the trolley deals with the threat that faced the five initially and we act because a side effect of turning the trolley (the one hit) will help us deal with a further bad side effect (organ failure) of what we do, an effect that threatens to undermine the good we have done for the five.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

So the bad side effect of the death of the one person helps sustain the good by defeating the undermining effect of organ failures. By contrast, in the first case, our goal of saving the five from *threats that face them initially* is only half achieved if the organ failure is also not stopped, and the death helps produce this new component of our initial goal.[25]

Richardson thinks there are worse problems with turning the trolley in the second case because we cause the organ failure problem. But would he think it was impermissible to turn the trolley away from people just because we thereby caused them organ failure even if this could be cured by another *non-bad side effect* of our turning the trolley, such as delivering a drug? Presumably not, so our causing the organ failure does not seem to play the inhibiting moral role that he thinks it does.

Some of the cases that Richardson discusses in dealing with the sustain/produce distinctions (and other distinctions) are wartime cases. Although I discuss war cases, I make a point in IE that some of the distinctions to which I attend may be of no ultimate importance in the context of war because certain factors override or nullify their importance. This is a particular aspect of the morality of war that I discuss elsewhere.[26] But let us consider an example that he discusses in any case. In order to criticize another distinction I draw between (i) a mere means to a great good and (ii) a component of a great good directly causing a lesser evil, Richardson discusses the following two cases: (1) In one case, a small part of a military facility (whose total destruction is conceived as a greater good) blows up. This causes as a side effect the deaths of some children and also a further fire that blows up the whole facility. (2) In the second case, we bomb an empty building[27] because the fire will keep parents away from rebuilding a military facility that is already destroyed, and the fire also kills children as a side effect. In both cases, only the non-functioning of the whole military facility would be a sufficient good effect to justify the deaths of the children as side effects. He says that in both cases, 'the bombardier's action is completed once the bombs have been released . . . Given the bombers' lack of ability further to intervene . . . I cannot see what moral difference it makes

[25] Richardson mistakenly describes my view about the cases a bit. Contrary to what he says, I do not think that in the first case, the greater good is wholly causally downstream from the death of the one. I think a component of the greater good – the five being saved from one threat – is present independently of the one being hit, but another component, saving them from the organ disease, is yet to be produced by this bad event.

[26] See my 'Failures of Just War Theory: Terror, Harm, and Justice', *Ethics* 114 (July 2004), pp. 650–92.

[27] Richardson actually speaks of bombing another military facility. But this could confuse the contrast between the cases, so I have substituted a non-military building.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

whether the evil* . . . is downstream of some partial realization of the greater good that would justify incurring it' (p. 86).

This explanation for why he thinks there is no moral difference between the cases would imply that there is also no moral difference between the following two cases: (a) we explode a bomb that moves a trolley away from five people thereby redirecting it to a track where the trolley kills one bystander, and (b) we explode a bomb that both causes the death of a bystander and also leads to the trolley being redirected from five people. In both these cases, too, the agent's action is completed once he sets the bomb. Does Richardson also think that there is no moral difference in these cases? Then he lacks a basic intuition on which much of the discussion of the trolley problem is based.[28] Perhaps he is not wrong in lacking this intuition, if the distinction that underlies it is of no significance. But I do have the intuition that my action in the second trolley case is impermissible and I am trying to locate the distinction that I think is responsible for it. As in the two wartime bombing cases, the first case of the set involves a greater good or part of it (or a means, such as turning a trolley, which has these as flip sides) causing the bad side effect. The second case in each pair involves a mere means to a greater good causing the bad side effect.

Richardson next discusses my direct/indirect side effect distinction. He says '[s]urely the presence of one more link in the causal chain is a difference lacking moral relevance' (p. 87). I have said that if it is we who introduce the additional link (as in a Rube Goldberg device), it makes no moral difference. I was suggesting, however, that if the difference in direct and indirect causal chains represents a difference between (i) something that we introduce causing harm and (ii) the environment into which we introduce something causing harm, this may, at least sometimes, give rise to intuitions of moral difference. Another set of cases meant to highlight this possible difference contrasts the following two cases: (1) We must use a gas to save five people, but the gas will kill an immovable bystander. (2) We must use a gas, in itself innocuous, to save five people, but the gas being in the air will displace the location of germs already present in the atmosphere so that they kill someone. I suggested that it would be permissible to proceed to achieve a greater good in the latter case but not the former.

3. Despite rejecting these distinctions that I emphasize, Richardson says that he endorses my emphasis on inviolability as

[28] It is interesting to note, however, that discussions of the Trolley Case tend to contrast it with the Bridge Case (where the person's fall is a means to stopping the trolley). They focus less on cases where a bomb that would stop the trolley kills a bystander. I think this is a mistake.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

non-subordination.[29] He says (p. 89): '[H]er account of inviolability as non-subordination centers on a certain type of asymmetry in the way parties are situated . . . that implies that the description of the subordinated person's position "makes essential reference to his usefulness to achieve a good for [the] other person" (165).' He thinks '[t]he core idea of this account of inviolability is actually quite simple. It hinges on whether or not someone's being harmed is a causal means to the greater good [that would alone justify our action]' (p. 89). I have already noted (p. 133) that Richardson's 'simple' view would rule out acting in Otsuka's Six Behind One Case, and in the Loop Case, among others, and I think this is wrong. As I said above, a simple view is no good if it does not explain things.[30]

While I think that Richardson's simple view is wrong, it is useful to note here how Norcross's analysis fails to even recognize the property of cases that are of concern to Richardson (and many others). This is an example of how Richardson draws finer distinctions than Norcross (though less fine than Otsuka). Consider Norcross's discussion of my Tractor Case in which a component of a good state of affairs is a flip side of turning the trolley away, namely the five are completely saved from the trolley threat. But they are still threatened by a tractor coming at them. Just sustaining what has been brought about so far will not save them. And there will not be any point in saving them from the trolley if the tractor will kill them shortly anyway. Stopping the tractor is produced in my version of the case by a bystander being gently pushed in front of the tractor by the diverted trolley, where he is killed by, but stops, the tractor. Norcross thinks that killing in the Tractor Case is the same as killing in the original Trolley Case. To support this view, he imagines a version of the Trolley Case in which after the bystander is killed, his head falls off. I agree this does not change the permissibility of turning the trolley. He then imagines his Heeeed Case, in which the head rolls into the tractor that is moving toward killing the five. He thinks that this case is also just like the original Trolley Case. In support of this, he says: : 'To claim that it is worse to divert in Heeeed than in the original trolley case would involve claiming that, at least

[29] Though he wonders why, if inviolability is a good, consequentialism cannot take it into account (p. 81). I see it as a status already had by persons, not something that is brought about, and I think of consequentialism as trying to bring about good states of affairs.

[30] Richardson has two complaints about my 'non-simple' view of inviolability as non-subordination. One is, essentially, that it excludes from the category of subordination many cases where harm to someone causes a greater good. A second complaint is that I 'stretch' (p. 89) my account to include too much in the category of subordination, as in my Trolley Tool Case. I am more sympathetic with the latter concern.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

in this context, it's OK to decapitate someone, so long as it doesn't do any good!' (p. 76).

I disagree with his analysis and I think he fails to distinguish the moral problems raised by the different cases. His supporting claim, that it cannot be wrong to do what you do just because it will be useful, is certainly true sometimes. Suppose the trolley, diverted from the five, killed one person whose head rolled in front of a tractor moving toward killing a different set of six people. It would still be permissible to turn the trolley. In this case, a death that is already completely justified by the five having been saved has a further effect – the head falling off and stopping an independent threat to others. By contrast, about the Heeeed Case, many will think as follows: I would not turn the trolley that will kill the one person, if I knew the five would die shortly of a tractor anyway. So, in this case, *it makes a difference to whether I am justified in deciding to do what kills someone* whether his being hit and its effect are useful. This is not true in the original trolley case or in the case where a different set of six people are saved from another threat. This is because, in these two cases, I would be justified in doing what kills the one bystander even if his being hit had no useful effect. I, Richardson and Otsuka, unlike Norcross, are aware of this difference, though we have different views about how it affects permissibility. Norcross does not even notice the difference.

Norcross also thinks that Tractor and Heeeed are morally the same. I agree, though I think they may both involve impermissible rather than permissible acts. However, I think that Heeeed can be misleading in a way that Tractor is not. I constructed the Tractor Case so that it was clear that the bystander's being hit by the tractor is what killed him. I did this in order to make it clear that his being hit (which causes death) – not something that happens after his death that will not make him any worse off, such as having his head fall off – is necessary to stop the tractor. By contrast in the Heeeed Case, when it is an effect subsequent to death (his head falling off) that is useful to stop the tractor, one may forget that the hit that leads to death is *required* in order that the subsequent useful head-popping occur. It is important to remember, when you decide whether to do what kills someone, that you are deciding on the basis of whether a hit that causes death will be useful, not just on the basis of whether something subsequent that does not make the person worse off will be useful. This is easier to keep in mind in my Tractor Case than in the Heeeed Case. His equation of the Trolley, Tractor and Heeeed Cases suggests that Norcross's failure to see distinctions does not mean that they are not present.[31]

[31] I shall return to the Tractor Case below.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

4. Let us return to Richardson's use of my claim about non-subordination. My own brief description of subordination, saying that it makes essential reference to someone's usefulness to achieve a good for another person, was probably much too brief. So, let me make clearer what I had intended. My way of understanding 'achieve' is limited to producing, and so it would distinguish between usefulness in producing and in sustaining. But even if someone's causal role need only be to sustain a greater good, this could also involve subordination. To see this, it is important to emphasize the role of the Rescue Test in my discussion. Suppose we turned the trolley in the Loop Case because the one would be hit and stop the looping. But now we see it would be possible easily to rescue the one person off the track. Refusing to do what would help him or giving the trolley an extra push so that it would keep him down, when this was not justified by some other property of the omission or extra push, would be in service of his sustaining rather than producing the greater good. Nevertheless, these behaviors would be wrong. The typical way of putting this is that the agent would then be intending the one's being hit. My non-state-of-mind way of putting it (in chapter 5 of IE) is that neither the omission nor the extra push has a non-causal, flip-side relation to a component of the greater good; they are extras, not needed to move the trolley away from the five, that are mere means to having the hit to the one person occur.[32]

But does the fact that we turn the trolley in the Loop Case because the one person's being hit will benefit the five put the person in the asymmetrical subordination relation to them? Is his position then describable as essentially making reference to the good of the five? Here it is a question of how we interpret 'essentially'. Norcross, unlike Richardson, thinks that subordination is involved in turning the trolley on the one in the original Trolley Case as well because doing so is only justifiable in order to bring about good for the five. But to my mind, there is a difference between (i) a position one occupies essentially making reference to others' good and (ii) the reason why one

[32] When Richardson attempts to summarize in his own words my Principle of Productive Purity, he says. 'One cannot justify harming someone, or involving him without his consent in a project that foreseeably will harm him, on the grounds that doing so is causally necessary to *producing* the greater good; however, one can justify doing these things on the grounds that doing so is causally necessary to *sustaining* the greater good' (p. 84). This is a misunderstanding. Not anything that causes harm to sustain a good is permissible. Richardson's misunderstanding may arise because he separates (1) The distinction between producing and sustaining the greater good, and (2) The distinction between being caused by a component of the greater good that is working itself out and being caused by something that is merely causally productive of the greater good. I emphasize that a bad effect that will occur as a result of a greater good working itself out may sustain that good rather than produce a further part of it. Hence, distinction (1) and (2) work together.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

occupies a position making reference to others' good. One can occupy a position that does not make reference to the good of others for other-regarding reasons. I thought that the Trolley Case was a pretty clear case of substitution rather than subordination, as I was using these terms. This is because the five being threatened and the one not being threatened are two sides of the same coin (i.e. non-causal flip sides). Therefore, the one being in a threatened *position* cannot be a position whose function it is to help the five (i.e. that essentially makes reference to the good of the five).

In many cases with which I am concerned and whose permissibility Richardson disputes, such as Loop, the bad that initially happens to the one comes as a result of an *initial substitution* (not subordination) of him. That is, in Loop, we send the trolley to him versus allowing it to go to the five. Of course, we do this only because the one's being hit will help prevent another bad effect for the five (the looping) of what we do. Or in my Secondary Transplant Case, we give scarce medicine to save five people from fatal disease A rather than to one different person instead. (This is substitution.) However, we do this only because we know that organs available from the one after his death will save the five from a fatal side effect of the medicine. I, unlike Richardson, take this to be permissible and different from a case like the original Transplant Case, where we deny someone a medicine (to which he would otherwise have a right), not in order (substitutively) to help a greater number of different people with the medicine, but in order that the death of the one person make available his organs for transplantation in the other people. It is only in the latter case that I would classify the position that the one is made to occupy as bypassing substitution initially and essentially involving subordination to the others.

The Extra Push and Rescue Cases (as I have conceived of them) are also important in this regard, for giving the extra push and not rescuing, if easy rescue becomes possible, would not be substitution but subordination of the one person. To see this consider the analogy with the Secondary Transplant Case: after having given the scarce drug to the five which produces the bad side effect we discover another easy cure for the disease that the one has, a cure which cannot help the five. If we denied the cure to the one person not because we give it to some other people where it would do more good but so that his organs are available for the five, he would be subordinated. The fact that nothing we do in giving the scarce medicine to the five rationally commits us to not subsequently rescuing the one person with the new cure supports the view that there is no essential subordination of him to the others in the Secondary Transplant Case.

It may be asked, however, whether the Tractor Case does not also involve an intermediate substitution, if we turn the trolley from five

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

to one? Not substitution that would be justified in the absence of the additional track going in front of the tractor. (Notice that this is different from 'in the absence of the person on the loop.') Here we can see how my account of the difference between sustaining the good in the Loop Case and producing the good in the Tractor Case illuminates (in a rather unexpected way) Judith Thomson's claim about the Loop Case that we cannot 'really suppose that the presence or absence of that extra bit of track makes a major moral difference' to whether we may turn the trolley.[33] If the additional piece of track going in front of the tractor, were not there and the trolley were the only threat, it would be permissible to turn the trolley, substituting one for the five. This should not change when the track is there, on my account of the Loop Case. And indeed, Thomson had meant that the additional piece of track would not change a permissible turning into an impermissible one. But now consider the Tractor Case. If the additional piece of track going in front of the tractor were not there and the five faced not only the trolley threat but rapid death due to the tractor as well, we should *not* turn the trolley, substituting the one for the five. My suggestion in IE is that when the additional track is there, our behavior should not change.[34] Here, the claim is that the extra bit of track does not change an impermissible turning into a permissible one.

Norcross's Moe Case (pp. 78–9) is, in some ways, structurally like my Tractor Case. In it, we must send up a rocket to stop a meteor from killing many people. Ultimately, for the rocket to go (1) an already dead person must be hurled against a switch and (2) also another person in the room must not be alive or else the triggering switch will not work. It turns out that what is needed to hurl the dead person will also, indirectly, kill the other person, thereby ensuring the second element that is needed in order that the rocket go. Norcross finds it incredible that it could be permissible to do what moves the dead man if killing the living man is a mere non-useful side effect, but it would be impermissible to do the same thing when the triggering device requires both (1) and (2). But the issue is that how a triggering device works could require us to treat people in different ways. For example, suppose that satisfying the 'demands' of the triggering device required our separately shooting the live person. Presumably, this would be impermissible (if the number of people to be saved is not above a threshold on constraints on harming, as Norcross posits).

[33] J. J. Thomson, 'The Trolley Problem', *The Yale Law Journal* 94.6 (May 1985), pp. 1395–1415.

[34] There are other cases, however, which suggest that sometimes the additional track can make a difference to what we should do. For an example, see the Component Case in IE.

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press

Still, I have to agree that the supposed moral distinction between the Loop and Tractor Cases plays an important role in the development of one of my principles of permissible harm, the Doctrine of Productive Purity (DPP). I explicitly say in IE that if there is no moral distinction between these cases and turning in the Loop Case is permissible, then the fact that harm to someone is required in the causal chain would never be the crucial element in making an act impermissible. What would be crucial is *how* the harm to the person comes about (an upstream event). For example, is the harm the result of a mere causal means to a greater good or of a greater good itself? By contrast, suppose that there is a moral difference between (a) harm being necessary *to produce* another element in an outcome that alone can justify that harm in the Tractor Case and (b) harm being necessary *to sustain* an outcome (by stopping a side effect of what we do from undermining the outcome) in the Loop Case. Then, even when the person comes to be harmed in the same way, a subsequent type of causal role for the harm (a downstream feature) could affect moral permissibility. While I do not think I am off my trolley, as Norcross says, it is possible that I should take a different turn on the trolley in dealing with the Loop and Tractor Cases. Then, the question would be whether to emphasize the DPP or (what we could call) the Doctrine of Bringing about Harm.

frances_kamm@ksg.harvard.edu

https://doi.org/10.1017/S0953820807002944 Published online by Cambridge University Press