# Off Her Trolley? Frances Kamm and the Metaphysics of Morality

ALASTAIR NORCROSS
*Rice University*

Frances Kamm's aptly titled *Intricate Ethics*[1] is a tour de force of what Peter Unger calls the 'preservationist' approach to ethical theory. Here is some of what she says about her methodology:

> Consider as many case-based judgments of yours as prove necessary. Do not ignore some case-based judgments, assuming they are errors, just because they conflict with simple or intuitively plausible principles that account for some subset of your case-based judgments. Work on the assumption that a different principle can account for all of the judgments. Be prepared to be surprised at what this principle is . . . . I say, consider your case-based judgments, rather than a survey of everyone's judgments. This is because I believe that much more is accomplished when one person considers her judgments and then tries to analyze and justify their grounds than if we do mere surveys. (5)

If I were to employ this methodology, I would barely produce enough material for a brief journal article, let alone a 500-page book. I could, of course, keep coming up with more examples to generate 'case-based judgments', but it becomes clear to me pretty quickly that I don't *need* many of them, though I do enjoy constructing them, and will gleefully do so a little later. I don't need many of them, because my judgments don't conflict with the simple intuitively plausible principle that accounts for any given subset of my case-based judgments. But enough about me, what of Kamm? Well, she apparently needs enormously many case-based judgments (henceforth simply judgments). In fact, to adapt a well-known proof to the effect that a certain philosopher has written an infinite number of books, here's a proof that Frances Kamm has constructed an infinite number of trolley cases. For any trolley case that Kamm has devised, there is a yet more intricate trolley case that she has also devised. My instinct, when confronted with so many cases, is to address each one in turn, what Kamm has to say about it, and explain why I disagree, when, as is most frequently the case, I do disagree. I have, of course, done that with my copy of Kamm's manuscript, which is now covered with mostly indecipherable scrawl. However, in the

[1] Frances Kamm, *Intricate Ethics* (Oxford, 2007). Unless otherwise noted, all page references are to this book.

© 2008 Cambridge University Press
doi:10.1017/S0953820807002919
*Utilitas* Vol. 20, No. 1, March 2008
Printed in the United Kingdom

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press

interests of producing an article that can be read in one sitting, without the aid of attention-enhancing drugs, I have been pretty selective in what I choose to talk about. What I will do, therefore, is to address the portions of the book that coincide most closely with my own interests, and attempt to give a flavor of the arguments. I won't have anything to say about her treatment of the views of other philosophers. I'm sure they are more than capable of furnishing their own replies. I will first address her methodology of placing great weight on judgments. I will then address her methodological claims about the killing/letting die distinction. Finally, I will discuss the core of her argument for a non-consequentialist principle of permissible harm. Both the subject matter and even some of the critical points I will make in this regard overlap considerably with Henry Richardson's contribution to this symposium, but my own position, of course, is far removed from his.

## METHODOLOGY

Consider Kamm's methodology, as quoted above. Her dismissal of 'mere surveys' of everyone's judgments is a little too hasty. Her reliance on her own intuitions, and the admonition to her readers to rely on their own intuitions, assumes a remarkably robust view of the moral epistemology of case-based judgments. My worry here is not just that I disagree with many, perhaps most, of her judgments. I am all too aware that non-consequentialists' intuitions diverge radically from my own. However, concerning many of Kamm's judgments, including some that seem to be crucial to her argument, I have found *no one* who agrees with her, and I have tried them out on many people, students, colleagues, friends, family, most of whose intuitions have not been tainted (or rather enlightened) by consequentialist reasoning. I will be discussing some of these examples later. The discovery that almost no-one, including fellow non-consequentialists, shares the judgments on cases that form important steps in the argument could, at least, give one pause.

Surveys of judgments can do more than simply reveal that one is in a very small minority, though. Properly conducted, they can challenge the reliability of using intuitions at all, at least intuitions of a certain type. A good example of this is provided by Peter Unger's 4-option trolley case.[2] Very roughly, here's the picture: An out-of-control trolley is heading down a track towards six people. Option 1 is to do nothing (nothing trolley-related at least) and let the trolley kill the six. Option 2 is to divert the trolley onto a side-track with three people, thereby killing the three. Option 3 is to divert a different runaway trolley

[2] Peter Unger, *Living High and Letting Die* (Oxford, 1996), pp. 88ff.

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press

containing two people into the path of the original trolley. This will stop the original trolley from doing further damage, but will kill the two. If the second trolley is not diverted into the path of the original trolley, the two passengers will be fine. Option 4 is to divert, by remote-control, a large man on motorized rollerskates into the path of the original trolley. This will stop the original trolley, but will kill the large man. Unger claims that, when presented with this case, most people respond that it is at least permissible to choose option 4, the rest choosing option 1. However, when presented with the case stripped of options 2 and 3, most people respond that it is not permissible to choose option 4 (renamed option 2), but rather that option 1 must be chosen. Furthermore, and here's the kicker, *everyone* says that the presence or absence of options 2 and 3 makes no difference to the permissibility of choosing option 4. I have presented this case to several classes. In each case I sent out half the students (randomly selected), and presented the 4 option case to the remaining half. In each case, a large majority (between 75 per cent per cent and 85 per cent per cent) said that option 4 is at least permissible, the rest said that only option 1 is permissible. All the students said that it would make no moral difference if options 2 and 3 were deleted from the case. I then brought in the other half and presented to them the two-option case formed by deleting options 2 and 3. In all but one case a significant majority (between 65 per cent and 75 per cent) said that only option 1 is permissible. In the other case, a slight majority said that option 4 is permissible. I then described options 2 and 3 and asked them whether they thought it would make a moral difference if they were available. They all said no.

My point in discussing this example is not to argue that Unger's claim that option 4 is at least permissible (in either case) is correct, though I obviously think it is, but rather to highlight the epistemic fragility of our intuitions in cases such as these. Even the classic pairing of the original trolley case with the transplant case can illustrate this. These cases are usually presented in the order Trolley–Transplant. Presented in this order (as I have done to many hundreds of students over the years), the overwhelming majority (more than 95 per cent) says that diverting the trolley is permissible. A slightly less overwhelming majority (about 85 per cent) then says that using the healthy man's organs to save the five is not permissible (I present the transplant case stripped of its institutional baggage, so as to avoid, as much as possible, intuitions about doctor–patient trust, and the possible bad effects on society of violating it). However, when I present the cases in the reverse order, a larger majority says that it is impermissible to do the transplants, and a much smaller majority says that it is permissible to divert the trolley. What these cases, and many more such examples, should make us consider is the possibility that our firmly held intuition about a

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press

particular case could just as easily have been radically different, if, perhaps, we had been asked to consider a different one first, or if the case had included other options that we judge (and would have judged) to be irrelevant, or perhaps even, if John Doris is to be believed,[3] we had discovered a dime in a coin return, or had had an extra cup of coffee for breakfast (I know Doris is talking about behavior, not judgment, but the possibility remains).

## KILLING AND LETTING DIE

Let's move on to the supposed distinction between killing and letting die. In *Intricate Ethics*, as in previous writings, Kamm argues against a common methodology, exemplified by Jim Rachels' famous Smith and Jones cases, and provides an argument to the effect that killing is worse, in itself, than letting die. Rachels' main argument consists of the following pair of cases:

> In the first, Smith stands to gain a large inheritance if anything should happen to his six-year-old cousin. One evening while the child is taking his bath, Smith sneaks into the bathroom and drowns the child, and then arranges things so that it will look like an accident.
>
> In the second, Jones also stands to gain if anything should happen to his six-year-old cousin. Like Smith, Jones sneaks in planning to drown the child in his bath. However, just as he enters the bathroom Jones sees the child slip and hit his head, and fall face down in the water. Jones is delighted; he stands by, ready to push the child's head back under if it is necessary, but it is not necessary. With only a little thrashing about, the child drowns all by himself, 'accidentally,' as Jones watches and does nothing.[4]

The two cases differ only in that Smith kills his cousin while Jones 'merely' lets his cousin die. If the difference between killing and letting die were morally significant in itself, claims Rachels, Smith's behavior would be morally worse than Jones' behavior. However, the two are equally reprehensible. So, killing is not worse, in itself, than letting die.

Rachels' argument follows a standard format for arguments about the moral status of certain factors: present two pieces of behavior that differ only with respect to the factor in question. If the behavior differs morally, the factor is morally significant in itself. If not, it is morally insignificant. This methodology is by no means uncontroversial, though. It rests on the assumption that, if a factor is morally significant in itself, it will be significant wherever it appears. Against this assumption, Kamm argues that a factor, such as the difference between

[3] See John Doris, *Lack of Character* (Cambridge, 2002).

[4] James Rachels, 'Active and Passive Euthanasia', reprinted in *Killing and Letting Die*, ed. Bonnie Steinbock and Alastair Norcross (New York, 1994), p. 115.

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press

killing and letting die, could sometimes be morally significant, and sometimes not. (The claim is not that other factors could sometimes cancel out the influence of the factor in question, but that the factor's influence could vary in cases where *all other factors are equal*.) 'A property can behave differently in one context than in another' (17). By way of explaining this, she gives her account, similar to her previous accounts, of the difference between killing and letting die. Here is the crucial part. According to Kamm, the following are 'definitional or essential properties of letting die, but not killing and, hence, are essential differences between the two' (19):

> 1. In killing, we introduce a threat that was not previously present; letting die involves, at least, a *potential threat already present*. 3 (ii) [I]n letting die, someone loses a life that he would have had only with our help at that time. 4. In killing, we initiate an interference with the victim; in letting die, we avoid being interfered with (by having to aid). (18–19)

I should note that these three properties, if they are to make out their case for being essential to letting die but not to killing, make at least implicit reference to a general doing/allowing distinction. In one obvious sense of 'potential', all killings involve at least a potential threat already present. If you *are* killed, it was true in advance that you *could be* killed, or *would be* killed if certain things happened. As for the threat being already present, presumably Kamm doesn't mean anything so silly as that there was some object in the physical vicinity that we could identify as 'the threat'. Likewise, all victims of killings would have required the 'help' of their killers not to lose their lives, just as in killing we avoid being interfered with (by having to refrain from killing). Intuitively, if we want to block the claim that all killings involve already present potential threats, we will have to exclude some killings on the grounds that the death would only occur if the agent *does* something. Similar tactics would be required to block the claims that all victims of killings would have required the help of their killers not to end up dead (not killing someone doesn't always count as helping?), and that all killings involve the killer avoiding being interfered with. This is not, in itself, a problem, but since these properties are supposed to be morally significant, the doing/allowing distinction on which they rest will also have to be morally significant. My worry, which I don't have space to explore here, is that there simply is no distinction in the vicinity of the doing/allowing distinction which is both metaphysically robust and clearly morally significant. This worry about metaphysical foundations applies also to many aspects of Kamm's proposed 'essence of non-consequentialist constraints on harming'.

So, how are properties 1, 3(ii), and 4 supposed to show that killing is worse in itself than letting die, and to explain how the factor could

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press

sometimes be morally significant and sometimes not? The crucial claim here is that some intrinsic properties of letting die are 'exportable' to some killings:

> If letting die has an intrinsic property that, when exported into a killing case, makes the case less bad, but killing has no property that has the same effect when exported into a letting die case, this would . . . be evidence for an intrinsic moral difference between killing and letting die. (412)

Or

> letting die would be morally better per se in virtue of that property. (19)

There are two important questions involved with this strategy. (i) Are the properties that Kamm identifies as essential to letting die but not to killing morally significant? (ii) If they are, should we conclude that letting die is 'morally better per se'? I won't consider (i) in detail here. I have already said that the supposed moral significance of these properties will have to depend on the supposed moral significance of the doing/allowing distinction, of which I am highly dubious. Furthermore, Kamm's arguments for the supposed moral significance of these properties depend on judgments that I simply don't share. So, let's move to (ii).

This is a difficult question to assess. Consider the following case for comparison: suppose that there are two classes of murder, first-degree and second-degree. Second-degree murder is less morally serious, let us suppose, than first-degree murder, because it lacks premeditation, which first-degree murder includes. Second-degree murder, therefore, has a definitional or essential property – lack of premeditation. Murder does not have this essential property. Some murders lack premeditation and some have it. Do murder and second-degree murder differ morally *per se*? The question appears to be ill formed. The correct comparison with second-degree murder in this case is not simply murder, but first-degree murder. The factor with moral significance is not the difference between second-degree murder and murder, but rather the difference between premeditated murder and unpremeditated murder. It is the presence or absence of premeditation that has significance. Likewise, when we consider Kamm's argument, the appropriate claim should be that the factors with moral significance are those that she claims to be 'definitional or essential properties of letting die, but not killing', not the difference between killing and letting die itself. For an example from the realm of prudence, consider my grandfather's claim that a pint of Guinness a day is good for you, better in fact than any other amount (including none). Suppose that you say to me that you have a glass of Guinness a day. I reply that I do better than you, because I have a pint of Guinness a day, and a pint of Guinness is better *per se*

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press

than a glass of Guinness, because a pint of Guinness is essentially a pint, but a glass may or may not be a pint. You'd probably look at me as though I were soft in the head! The point here, of course, is that, as regards Guinness, it's the difference between a pint and, say, half a pint that has prudential significance, not the difference between a pint and a glass.

What of Kamm's claim that 'a property can behave differently in one context than in another'? This is supposed to support the view that, say, the difference between killing and letting die is sometimes morally significant and sometimes not. But this way of looking at it assumes that we can clearly separate a property from its context, at least in those cases in which the context is claimed to make a difference. But consider Rachels' examples again. One of the common objections to it is that it may not be possible to generalize from it to the kinds of cases that most concern us, that is, cases of euthanasia. The example involves two pieces of malevolent behavior. Perhaps there is no moral difference between malevolently killing and malevolently letting die. But we can't necessarily generalize from this to cases of euthanasia, which involve benevolent motivations. Perhaps benevolent killing *is* worse than benevolent letting die. Of course, this kind of objection can always be raised against an argument from a matched pair of examples. Whatever the distinction we are trying to isolate, if the examples differ only with respect to that distinction, there will be any number of other factors held equal that could have been different. Merely identifying one such factor poses no real threat to the original argument. Rachels' examples are both in bathrooms, but it would be ridiculous to suggest that we couldn't generalize to killings and lettings die in other locations. It might be objected that the difference between malevolent and benevolent motivation is intuitively more morally relevant than the distinction between one location and another. Maybe so, but what is intuitively (though not uncontroversially) relevant is the distinction between a piece of benevolent behavior and a piece of malevolent behavior. It is not intuitively obvious that the distinction between a distinction involving malevolent behavior and a distinction involving benevolent behavior is morally relevant. At the very least, the objector to Rachels' argument owes an explanation of why such a distinction should make a moral difference between distinctions.

The preceding discussion illustrates some important methodological points concerning arguments for the moral relevance or irrelevance of distinctions. It is always possible to object to an argument from a pair of examples that a particular distinction is sometimes morally relevant and sometimes not. The objection by itself, though, has no force. It must be accompanied by an explanation of why the distinction is morally relevant in some cases and not in others. This is much harder

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press

to provide. At the very least, then, Rachels' argument puts the burden of proof squarely on those who would invest the distinction between killing and letting die with moral significance.

## NON-CONSEQUENTIALIST CONSTRAINTS ON HARMING

Chapter 5, 'Towards the Essence of Non-consequentialist Constraints on Harming', is where a lot of the action is. Trolleys, tractors, bombs and gases proliferate madly. Victims die under the wheels of trolleys or tractors, under piles of falling stones, or collapsing roofs; they are gassed or splashed with acid, and some of them are just plain blown up. The shared feature of the examples is that they involve a greater good and a lesser evil, as for example, the greater good of saving five and the lesser evil of killing one in both the original trolley case and the transplant case. Of course, this being a non-consequentialist theory, sometimes it is permissible to act so as to bring about the greater good at the expense of the lesser evil, and sometimes it is not. All of this mayhem is aimed at working out a non-consequentialist Principle of Permissible Harm (PPH), part of which is a Doctrine of Productive Purity (DPP). 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. A lot of rather metaphysically suspect distinctions play important roles. It apparently can matter whether the lesser evil is *directly* or *indirectly* caused, by, for example, turning the trolley. It can also matter whether the greater good is *produced* or *sustained*. Especially important is whether the means that leads to the lesser evil also *causes* the greater good, or rather has the greater good as what Kamm calls its *non-causal flip side*. In case you're wondering, it's considerably better for it to be the non-causal flip side (don't ask me why). It is also supposedly important to individuate threats quite finely. So, in the loop version of the trolley case (in which the diverted trolley would loop around and kill the five, if the one person weren't on the track), the initial threat to the five is being hit by the trolley from one direction, and only when the trolley is diverted onto the looping side track does another threat arise, that of being hit by the trolley from a different direction. Although the principles involved are, quite literally, incredibly complicated, they supposedly express a fairly simple concept, that of the 'inviolability' of persons. The crux, or at least one of the cruces, of this concept involves another distinction, that between people being *substituted* for others and being *subordinated* to others. Again, in case you're wondering, substitution is more likely to be permissible than subordination. Much as I would like to work through each step of the argument, as I said above, space

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press

is limited, so I will try to highlight some of what appear to be the most important moves, and the most implausible claims.

Consider the loop version of the trolley case. As Henry Richardson says (in his contribution to this symposium), Kamm devotes a considerable amount of time to it and its variations. Richardson thinks that she goes wrong at the very beginning in agreeing with Thomson's claim that it is permissible to divert the trolley in this case. If only she agreed with him, that diversion is impermissible, she could give a nice simple account of inviolability, involving a (possibly defeasible) prohibition against causing harm to some as a means to benefiting others. For once, I find myself in agreement with Kamm's (and Thomson's) intuitions on a case here. (I will briefly consider Richardson's suggestion later.) So, given that Kamm is obviously not content with the simplest explanation of why it is permissible to divert in the loop case, what is *her* explanation? In this regard it will help to consider her explanation of why it is permissible to divert in the original trolley case. In this case, the means that causes the lesser evil has the greater good as its non-causal flip side. This is in contrast to the transplant case, in which the means that causes the lesser evil also *causes* the greater good. What, exactly, is a non-causal flip side? Here is some of what Kamm has to say in explanation of this idea:

> not all means to a greater good *cause* that greater good, as they can have a noncausal relation to it . . . . In the context where only the trolley threatens the five, the redirection of the trolley threat away from them – by which I mean the moving of the trolley itself – is a means of saving them. It is the same event as their becoming free of threats, and this is the same event as their becoming saved. Put another way, in the context of the original Trolley Case, the five being free from threats is constituted by the trolley being away from them. Hence, there is a noncausal relation between the trolley turning away and the five being saved. (140–1)

The first thing to note about this is that, on some widely endorsed accounts of causation, turning the trolley away from the five straightforwardly causes the five to be saved. So it appears that Kamm's approach at this point requires a particular account, or at least a type of account, of causation. The second point is that the appeal to context is crucial here. We are only inclined to agree (if we are) that the five being free of threats is constituted by the trolley being turned away from them in this context, because we understand the context to include certain physical guarantees that the trolley won't come back to hit the five. In this case, the guarantees are provided by the fact that the tracks don't come back, and that trolleys don't jump from their tracks and back onto their original tracks. But what about the transplant case? What if the doctor is guaranteed to transplant the organs into the five once he removes them from the one? Why then couldn't we say

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press

that, in *this* context, the five being free of the threat of organ failure is constituted by the removal of the organs from the one? We could, of course, insist on the fact that the doctor couldn't be *guaranteed* to complete successful transplants, given the number of things that could go wrong, and his own power of choice. Quite apart from the unseemly intrusion of mundane realism into our fantastical cases, this doesn't address the possible variation of the transplant case in which the whole process is carried out by a machine, which can't be stopped after the doctor presses the 'Start Transplants' button.

Let's pretend for the sake of argument that the notion of a non-causal flip side will both stand up to critical scrutiny and distinguish between Trolley and Transplant in the appropriate way. How does this apply to the Loop Trolley case? Surely, in this case, the diversion of the trolley can't have the greater good as its non-causal flip side, because merely diverting the trolley doesn't extinguish the threat of the trolley hitting the five. Here's where some fine-grained threat individuation is needed. Apparently, diverting the trolley in the loop case does have as its non-causal flip side 'the five people being rid of the first problem involving the trolley' (154), that is the problem of being hit from one direction. There is of course the inconvenient fact that there is also the problem of being hit from a *different* direction, but we can say that that is a *new* problem that only arises as a result of the elimination of the first problem. This allows us to say that the lesser evil doesn't actually *produce* the greater good, but rather only *sustains* it, by defeating defeaters to it. Furthermore, Kamm claims that these considerations allow us to say that, 'in redirecting the trolley, we have achieved a state that, when we abstract from further consequences for the five of moving the trolley from its initial course, would be the greater good. I shall call such a state the "structural equivalent of the greater good"' (154). The fact that the means that actually stops the trolley from hitting the five, by diverting it into the one, has the structural equivalent of the greater good as its non-causal flip side helps to ensure its 'productive purity', and thus to render it permissible.

This last sequence of moves illustrates the Ptolemaic character of Kamm's employment of her methodology. New crystalline spheres and epicycles are constantly being added in an attempt to fix the appearances. Many times, the additions can appear *ad hoc*. Was the problem facing the five really just the problem of the trolley heading towards them from *that* direction? If so, why wasn't it the problem of the trolley heading towards them at that *velocity*? In which case, what would Kamm say to the following variation? A trolley is headed towards five. There are no side tracks, but there is a very large person walking towards the track, about to cross it. At the present velocity of the trolley, the person will cross the track before the trolley reaches that

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press

point. However, we can speed up the trolley enough to hit the person, thereby stopping the trolley and saving the five, but killing the person. Speeding up the trolley has as its non-causal flip side the five being rid of the problem of being hit at *that* velocity. However, it also gives rise to the new problem of the five being hit at a different velocity. Does the lesser evil of the one being hit in this case *produce* the greater good, or just *sustain* it? When we abstract from the further consequences of speeding up the trolley – that the five will now be hit at a different velocity – speeding up the trolley has the greater good as its non-causal flip side. Therefore, speeding up the trolley actually has the structural equivalent of the greater good as its non-causal flip side. I have no idea what Kamm would say about this case. Inasmuch as I understand how the minds of non-consequentialists work, I suspect that most of them would object to this speeding up. But, as I said before, Kamm has intuitions that seem to diverge from those of most of the non-consequentialists I know. To return to the original loop case, if I had to say what is the problem facing the five, it would be the problem of being hit (and killed) by the trolley, not being hit from this direction or that.

Consider now the tractor case with which Kamm contrasts the loop case:

> The trolley is headed toward killing the five. It can be redirected and there is no way for it to loop back. However, there is another threat, a deadly tractor, also headed toward killing the five shortly . . . . We know that if we redirect the trolley, it will gently hit and push (without hurting) one person into the path of the tractor. His being hit by the tractor stops it but also kills him. (137)

Kamm claims that it is impermissible to redirect in this case. The redirection doesn't have the elimination of all the problems that were previously facing the five as its non-causal flip side, because, even though it has the elimination of the problem of being hit by the trolley as a non-causal flip side, it is *causally* related to the elimination of the problem of being hit by the tractor. What should we think about this case? First, let's add a detail to the original trolley case that is consistent with how it is usually described. When the trolley hits and kills the one, it severs his head, which rolls down a gentle slope. That detail has to be morally irrelevant. No one can honestly say that they agreed with the permissibility of redirecting in the original trolley case only because they assumed that the one's head didn't pop off. Now consider a variation of the tractor case, that I call 'Heeeed' (this is a reference to the movie 'So I married an Axe Murderer', in which Mike Myers plays both a young man and his cantankerous Scottish father). Just as in the original trolley case, the one is on the side track, and will be killed if the trolley is diverted. Furthermore, he has an unusually large head, which will be severed by the trolley and roll down a gentle hill. Just as in the tractor case, there

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press

is a deadly tractor headed towards killing the five. If the head rolls down the hill, it will stop the tractor. Now, it seems to me that Heeeed is morally equivalent either to the original trolley case, or to the tractor case. Furthermore, it seems clear to me that Heeeed is morally equivalent to *both* the original trolley case and the tractor case. Consider them in reverse order. Can we honestly say that it makes a difference whether the trolley pushes the one into the path of the tractor, which then kills him, or pushes his head into the path of the tractor, *after* killing him? We can even specify that, in the original tractor case, the tractor decapitates the one, to equalize the gruesomeness. Well, what about the relation between Heeeed and the original trolley case, with the decapitation detail spelled out? To claim that it is worse to divert in Heeeed than in the original trolley case would involve claiming that, at least in this context, it's OK to decapitate someone, so long as it doesn't do any good! Imagine that we are just about to permissibly divert a trolley that will decapitate someone, when a big tractor hoves into view. Now, all of a sudden, it's impermissible. Something is clearly amiss here. This also raises a problem for Henry Richardson's simple approach that focuses on whether the lesser evil is a means to the greater good.

Let's move on to a pair of examples that, more than any other, illustrate Kamm's proclivity for concocting dubious distinctions and investing them with moral significance. The first example is a three-track version of the original trolley case. The trolley is headed towards five, but there are two other tracks, onto either of which we can redirect the trolley. On one track there are three people, and on the other there is one person. Clearly here we can redirect to the track with one, and Kamm agrees. Now consider a slight variation, that Kamm calls a 'tree trolley' case:

> Suppose that the trolley that is headed toward three people is a trolley we, as bystander agents, have already redirected away from five people. En route to the three, it can be redirected again to a track where one person will be hit. (It was not possible to redirect it to that track initially.) (144)

Obviously, in this tree trolley case it is permissible to redirect from five to three (Kamm agrees). The question is whether it is then permissible to redirect a second time from three to one. Now, to everyone to whom I have presented this case (many of whom are unrepentant non-consequentialists) it is a no-brainer. *Of course* it is permissible to redirect a second time. However, as you have probably guessed, I wouldn't be discussing this case if that were also Kamm's answer. You might be wondering at this point whether the wheels have come off the trolley, but bear with me. What does Kamm think is the relevant difference between this case and the first three-track case? Well, in the first case, before the redirection, the three and the one were all

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press

'physically susceptible to harm'. In the second case, before the first redirection, only the three were physically susceptible to harm. The one only becomes susceptible as a result of the first redirection. So, why wasn't the one originally physically susceptible to harm? Because he was 'not reachable by means of a single redirection' (145). Not being originally physically susceptible to harm doesn't guarantee that one can't be permissibly harmed by multiple redirections (Kamm has another example to illustrate this claim), but it is supposed to make the difference here. But wait a minute. What is meant by a 'single redirection'? Presumably, it can't be just a matter of a single *decision*. I could very easily decide at the beginning that I will pull the first lever (or press the first button) and then immediately pull the other one. Does she mean a single physical *action* (whatever that ultimately is supposed to amount to)? But what if the two sets of points were only a few feet apart, and it were possible to connect the two levers so that pulling the lever to redirect from five to three would also pull the second lever? Or what if there were two buttons we could press, one of which would switch the track from five to three and the other would switch both tracks simultaneously? Wouldn't we then be able to perform a single redirection to get the trolley both from the five to the three and then from the three to the one? If these possibilities would count as single redirections, the moral difference could be made by the possibility of tying two levers together! Perhaps, though, Kamm would claim that these wouldn't be single redirections, because, even though they only require a single *action*, the action leads to two different *events* (two sets of points switching). But, for this to work out exactly as she wants it to is going to require some pretty hefty voodoo metaphysics, which I don't have the space to explore here.

This tree trolley case is also supposed to illustrate the distinction, which both Kamm and Richardson believe to be crucial to the notion of inviolability, between *substitution* and *subordination*. In the first three-track case, the three and the one are all *substitutable* for the five. But in the tree case, the three are *substitutable* for the five, but the one is *subordinated* to the three. It's not really clear why this is supposed to follow, but she has a more detailed account of this distinction a little later:

> When we do something bad to someone as a causal means to save someone else from a threat, the first person occupies a different position than the second person would have in being threatened. The position he occupies involves subordination to that other person, because the position he occupies – as the means to the good of another – makes essential reference to his usefulness to achieving a good for that other person (165).

I have a hard time seeing how this account will do the requisite moral heavy lifting by itself. Consider the permissible case of *substitution* in

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press

the original trolley case. The position occupied by the one in this case – as someone whose death is an acceptable side-effect – makes essential reference to achieving a good for the five. That is, it is only morally justifiable to bring about the death of the one, because in doing so we achieve a good for the five. So we don't here have any enlightenment as to the difference between substitution and subordination. Kamm attempts to spell it out in yet more detail with a variant of the trolley case, the Trolley Tool Case. The trolley is heading towards destroying a tool, but can be diverted to a track with one person on it. Ordinarily, we couldn't justify killing the one to save a tool, but what if the tool itself could be used to save many lives (and presumably nothing else could save these lives)? Even here, she claims, we can't redirect. Why not?

> This [would imply] that the many people are transmitting their importance to the tool; these people would have the power to change the value of the things in the world we all share, so that the nondestruction of the one person is less favored than the nondestruction of the tool. This elevation of the tool over a person, in virtue of its usefulness for others, results in a subordination of the one person to the other people, I believe. Hence, subordination is a broader category than using someone as a mere means. (166)

But this is confused. Inasmuch as the many are 'transmitting their importance to the tool', so is the one. After all, even if the tool couldn't be used to save others, but only, let's say, to bring some enjoyment, we would redirect, *if the one weren't on the track*. His presence on the track in this case would make it impermissible to redirect. So he would be transmitting his importance to the tool by changing its value negatively. In Kamm's case, the many transmit more positive importance to the tool than the one transmits negative importance. It's just his bad luck that they, collectively, have more importance than he does. It doesn't change the fact that, if any transmitting gets done, he does his fair share.

The problem I have with Kamm's approach, then, as illustrated by these brief snippets, is that the whole intricate edifice, besides being based on intuitive judgments that not only I, but many non-consequentialists, find incredible, also rests on such shaky metaphysical foundations that the slightest ontological breeze could bring it crashing down.

So why not adopt the more traditional non-consequentialist suggestion of a constraint against using as a means? Apart from the unpleasantness of having to ban redirection in the loop case, consider the following case: A meteor is heading towards destroying many people (for those who espouse constraints with thresholds, set the number just below the relevant threshold). We devise a missile that will divert it harmlessly, if launched at the right time. The launch mechanism is risky, so only two qualified scientists, Homer and Barney, are allowed in the launch room. The procedure involves pushing the launch button

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press

at noon on the relevant day, and then vacating the room, before the launch takes place. If the button isn't pressed within ten seconds of noon, the meteor will not be destroyed, and the many will die. Shortly before the scheduled launch, there is an accident in the launch room, which kills Barney and renders Homer unconscious. General Moe, outside the room, sees what has happened. He can't get in, but he can break a window, which he does. Barney's dead body is slumped in front of the button. Moe has a grenade, which he can toss into the room. The explosion will push Barney's body onto the button, launching the missile, but it will also kill Homer. In this case, Homer's death is not a means to saving the many. (Perhaps Kamm would object to the grenade tossing, because it directly causes Homer's death. Suppose, then, that the blast from the grenade doesn't kill Homer directly, but creates vibrations which cause part of the roof to fall and kill him. This variation demonstrates the moral irrelevance of the direct/indirect distinction.) Suppose, then, that Moe has acted permissibly, because he hasn't used Homer as a means. But what if he discovers afterwards that the triggering device had been recently fitted with a safety mechanism, which would only allow the launch to proceed if no signs of human life were detected inside the room? Now, it turns out that Moe's action is impermissible, because Homer's death is a means to the saving of the many. But this would make the permissibility of saving many depend on the mechanics of the triggering device.

This last example illustrates a problem that is shared by Kamm's labyrinthine non-consequentialist constraints and by the much simpler non-consequentialist constraint against using as a means. In these non-consequentialist approaches, concern for people (and other moral patients), rightly expressed as concern with how they fare, is, at least sometimes, replaced (or at least overridden) by concern with the causal mechanics of how they fare. The consequentialist is concerned with the harms and benefits that befall the moral community (agents and patients). So is the non-consequentialist, but she is also concerned with such details as whether particular harms and benefits are causally 'upstream' or 'downstream' from other harms and benefits, and some, such as Kamm, are even concerned with how many links there are in the causal chains connecting such harms and benefits to each other or to an agent's actions. Of course consequentialists are concerned with means as well as ends, but their concern is not with means for their own sake. If one means to a particular end contains more suffering than another, the consequentialist prefers the latter, because she cares about suffering, not the details of the causal route. The non-consequentialist will, perhaps, claim that it is 'respect' for persons, or the 'inviolable' status of persons, that prohibits certain causal routes and permits others. Thus, adhering to certain non-consequentialist

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press

constraints simply *is* an expression of concern for persons. But, if the notions of respect or inviolability, at least as understood by non-consequentialists, really just amount to constraints on the permitted causal routes to and between harms and benefits, it is hard to see them as anything other than fetishizing certain causal processes. It is the sort of approach that could make the mechanics of the triggering device relevant to the permissibility of saving many at the cost of one. How could this be relevant to justified moral concern for others? As I said, the non-consequentialist can try to cover up her fetishizing preferred causal mechanisms with talk of respect or inviolability, but, as Abraham Lincoln said in response to his own question, 'If you call a tail a "leg", how many legs has a dog?', 'calling a tail a "leg" don't make it a leg'. I admit I don't have a knock-down argument against such fetishizing. I find the bureaucratic obsession with the details of the causal routes (sometimes at the expense of the well-being of real people) morally alien. It is the accountant's obsession with dotting every 'i' and crossing every 't'. It is the jurist's obsession with following the rules, which can even rule out of court evidence of a defendant's innocence, on the grounds that the paperwork was filed incorrectly. Of course, this might simply be a matter of philosophical taste, and so this piece of psychological autobiography might do no more than explain why I am a consequentialist. I suspect, however, that it goes deeper than that, but this is not the place to explore it further.

The previous paragraph doesn't address one important source of non-consequentialist constraints. Some non-consequentialists (though not Kamm) appeal to the supposed significance of an agent's intentions to justify non-consequentialist constraints on harming. Such a move has its own, in my view insurmountable, problems, but that is a topic for another paper.

In conclusion, if we can justify neither Kamm's complex constraints nor the simple prohibition against using as a means, how are we to proceed in the quest for constraints on harming? I would humbly suggest that we simply drop the prefix from 'non-consequentialist', but I doubt that that suggestion will find favor among my fellow symposiasts.

norcross@rice.edu

https://doi.org/10.1017/S0953820807002919 Published online by Cambridge University Press