

Responsibility, Permissibility, and Vicarious Agency

Jeff McMahan

Intricate Ethics is not only brilliant but also vast, so that a short commentary can discuss only a tiny proportion of its contents. I will focus on the central claims of the chapter called “Responsibility and Collaboration.” These are:

(1) When one person acts as the vicarious agent of another, there can be conditions in which responsibility for the bad effects of the agent’s act transfers entirely to the principal, leaving the agent with *no* responsibility for those effects.

(2) The location of responsibility for the bad effects of an act can make a difference to the permissibility of the act. Whether responsibility lies partly with the agent or entirely with the principal can make the difference between the permissibility and impermissibility of the agent’s act.

Kamm takes as her starting point the well-known example from an early essay by Bernard Williams in which a Captain in a South American army is about to execute twenty innocent Indians when a traveler, Jim, appears. The Captain, in a gesture of magnanimity, offers to free nineteen of the Indians if Jim will shoot one, whereupon all the Indians beg Jim to accept this offer. Kamm observes that many factors seem to contribute to the permissibility of Jim’s killing one: that it would be in the *ex ante* interest of all the Indians, that it would not be worse for the one whom Jim would kill, who would also have consented to be killed, and so on. Kamm’s concern, however, is with another

factor that she thinks is important to the permissibility of Jim's killing an Indian: namely, that the initiative for the killing comes from the Captain's offer rather than from Jim.

How might Jim have acted on his own initiative? There follows one of the hypothetical cases for which Kamm will be forever celebrated in song and legend. She imagines Jim observing the proceedings from a distance, equipped with an infallible telekinetic brain scanner that reveals the Captain's determination to release nineteen if someone else kills one first. Kamm claims that if Jim kills an Indian in this case ("Scan"), *some* of the responsibility for the killing will remain with him. But in a parallel case ("Offer") in which the Indians also have no opportunity to consent but in which the Captain offers Jim the option of saving nineteen by killing one, she claims that full and complete responsibility for the death transfers to the Captain, making it "more permissible" for Jim to kill in this case than in Scan – if indeed killing one is permissible at all in Scan.ⁱ She concludes that on some occasions "collaborating with evil" to bring about a good outcome can be morally better than acting on one's own.

The case Kamm makes for these claims relies largely on intuitions about these and other examples. She suggests that "characterizing Jim as a *substitute actor* who acts on behalf of someone else comes closest to capturing the way in which responsibility gets shifted" in cases such as Offer but not Scan. Using the capitalized term "Agent" to refer to a substitute actor, Kamm cites a further example of substitute agency that is presumably intended to strengthen her case:

A lawyer who carries out an eviction of a poor tenant is the Agent of his client who owns the building. When it is permissible for the lawyer to be an Agent in this capacity, he is not morally responsible for the bad effects

of his act; it is his client who is morally responsible and accountable. Responsibility gets shifted. Indeed, it can sometimes be morally permissible to be an Agent and do an act as an Agent, without its being morally correct for anyone (including the client) to do the act if one is not an Agent. So insofar as a client has a right to do what is wrong, he may sometimes do something any person morally ought not [to] do through an Agent who acts permissibly qua Agent. (312)

This example is of a sort that is more common than any of the cases involving Jim, who acts as an Agent in a way that has bad effects *only* in order to avert even worse effects. I will return to it later.

For the moment, let us consider the plausibility of Kamm's two main claims, beginning with the claim that it makes a difference to the location of responsibility for Jim's act of killing whether the impetus for it comes from Jim or the Captain. Kamm says that in Offer "the Captain has responsibility for creating the contingency of the better outcome," while in Scan it is Jim who "has responsibility for initiating" that contingency. (311) Note, however, that the claim that the Captain bears sole responsibility for the killing in Offer is an *inference* from, rather than a *restatement* of, the claim that he has responsibility in that case for making that killing a means of saving nineteen. Indeed, "responsibility" has different meanings in these two claims. When Kamm says that in Offer the Captain has responsibility for *creating the option* involving fewer killings, she seems to mean only that this option is a foreseeable and intended consequence of his voluntary action. But when she says he is solely responsible for the *killing*, she means two things: that he is "completely morally responsible in the sense of

being to blame for, being at fault for, the negative consequences of Jim's act," and that he is "completely responsible, in the sense of being accountable... (i.e., liable for criticism, punishment, or compensation for the death)." (311) In the same paragraph, she uses "responsibility" in two further senses: causal responsibility and "capacity-responsibility." Neither of these, however, is the kind of responsibility that she thinks gets shifted in Offer, for Jim clearly causes the death of the Indian he kills and does so as an agent capable of morally responsible action. The responsibility that she believes is transferred from Jim to the Captain is a matter of blameworthiness and liability.

Kamm provides no principle that governs the transfer of responsibility in cases such as Offer. Instead she simply sets out the contrast between Offer, in which we are supposed to see that responsibility fully transfers, and Scan, in which it supposedly does not. Yet there are many possibilities between these two cases. Suppose, for example, that Jim has done the scan and is muttering to himself about whether he should kill an Indian. The Captain overhears Jim debating with himself and this prompts him to offer to release nineteen if Jim will kill one. Otherwise he would not have made the offer. Or suppose that Jim has done the scan and has thus learned three facts about the Captain's mental states: that he is disposed to release nineteen if someone else will kill one, that he does not intend to offer anyone the option of killing one as a means of saving nineteen, and that he enjoys forcing people to act in ways to which they are morally averse. So Jim remarks to the Captain, "It certainly would be awkward for someone if you were to offer to release nineteen of these Indians if he would kill one," and this then prompts the Captain to make Jim the offer. Or suppose that Jim has done the scan and the Captain has made no offer but says to himself, within Jim's hearing, "If only someone would

shoot one of these Indians, I would release the others” – rather in the manner of the king’s exclaiming of Thomas à Becket, in the presence of several knights, “Will no one rid me of this turbulent priest?” If Kamm is right that who is responsible for the killing depends on who is responsible for initiating the sequence whereby killing the one becomes a means of saving nineteen, we need a criterion for locating the latter form of responsibility in cases such as these. We need to know the precise grounds on which blameworthiness and liability are supposed to transfer from Jim to the Captain, and from Agents to principals more generally. All Kamm says is that it is a matter of “the *initiation* of a lethal plot.” (312) But in the cases just cited, both agents share in or contribute to the initiation, each acting in ways that may or not be relevant to the location of responsibility. I suspect that even if one could find a sharp criterion that would capture the spirit of Kamm’s proposal and make fine discriminations among the kinds of case just noted, its assignments of blameworthiness and liability would seem intuitively ad hoc.

Suppose that in Scan Jim is a foreign spy, so that he cannot reveal his presence, thereby giving the Captain an opportunity to make him the offer. My intuition is that in these conditions, Jim is not only permitted but morally required to kill one of the Indians. If that is so and if, acting for laudable motives and intending only to save as many Indians as he can, he does indeed kill one, I can see no grounds for blaming him or for holding him liable to defensive action, punishment, or payment of compensation to the victim’s relatives (who deserve compensation, just not from Jim). In this case as in Offer, all responsibility, in the relevant senses, belongs to the Captain.

But suppose that, unlike me but like Kamm, one has the intuition that some, or perhaps all, of the responsibility for the killing remains with Jim. What follows? Suppose that, as in the original version of Scan, Jim is not a spy. And suppose he has only a cheap brain scanner that has indicated with certainty that the Captain will release nineteen if someone else kills one now, but is not sufficiently accurate to determine with certainty whether he would make an offer if Jim were to approach him. It can detect only that the probability of an offer is high. It does, however, reveal a mounting irritation in the Captain, and predicts, with a high degree of reliability, that his conditional intentions, and the terms of any offer he might make, will become less generous as he grows more irritable. If, moreover, Jim decides to approach the Captain, he must do so unarmed, thereby losing the option of killing an Indian except by invitation, or by offering to the Captain to kill one in exchange for the release of nineteen, in which case, as Kamm observes, he will initiate the lethal sequence and thus be responsible in much the way he would be if he were simply to shoot now.

On Kamm's view, if Jim wishes to avoid any responsibility for the killing of an Indian, he has a reason not to kill on his own initiative but to approach the Captain instead. For only by approaching the Captain does he have a chance of receiving an offer and thus being able to save nineteen without incurring any responsibility for a killing (assuming that merely approaching the Captain does not count as initiating a "lethal plot"). Yet if he decides to approach the Captain, there will be a significant risk that his options will be reduced to these: (1) saving eighteen or fewer Indians by killing two or more in response to an offer and thus without incurring any responsibility, and (2) offering to kill two or more in order to save eighteen or fewer, thereby incurring

responsibility much as he would if he were to kill one now, without approaching the Captain. Of course, his reason to risk having to kill more Indians is not decisive. But there is, on Kamm's view, *a* reason, so that if the probability of Jim's receiving an immediate offer from the Captain is high enough, the reason to approach the Captain will become decisive, even though it is certain he could save nineteen if he were to kill one now and there is some risk that by approaching the Captain his best option may be to kill two or more to save the remainder. I find it hard to believe that there is any reason to take that risk.

It is also hard to believe that it is in the Captain's power to determine where the blame for the killing will lie. Suppose the Captain knows that Jim has done the scan but also knows that Jim is unaware that he has this knowledge. As Jim approaches, the Captain considers whether to offer the option of killing an Indian or to hold out until Jim feels compelled to offer to do the killing. Can the Captain's decision really determine whether the blame for the killing will lie only with him or whether he will share it with Jim? If it can and the Captain wants to avoid complete responsibility for the killing, he had better not make Jim the offer (though sharing the responsibility with Jim would not *entail* a reduction of his responsibility).

Consider now Kamm's second main claim: that the permissibility of Jim's killing an Indian is affected by whether some or all of the responsibility will lie with him or whether complete responsibility will lie with the Captain. Kamm's view that the permissibility of an act can depend on who will have responsibility for its bad effects is widely shared. In his account of his well-known experiments on obedience to authority, Stanley Milgram describes how some subjects were reluctant to act in ways they believed

would seriously harm innocent people but overcame that reluctance when told, in response to their own questioning, that responsibility for the harm would lie with the experimenters rather than with them.ⁱⁱ Military commanders sometimes persuade themselves of the permissibility of attacking a civilian area on the ground that responsibility for the deaths of civilians will belong to those who initially fired weapons from within that area. Similarly, Alan Dershowitz has proposed that, to deter terrorist attacks, Israel should designate in advance a Palestinian village “used as a base for terrorists” that will be destroyed in the event of a Palestinian terrorist attack, and claims as justification for this strategy that if Israel fulfills such a threat, all responsibility will lie with the terrorists: “The destruction is entirely their own fault.”ⁱⁱⁱ Such beliefs are common but it is difficult to avoid the suspicion, in these instances at least, that they involve an element of bad faith, a clutching at straws of exculpation. (Why, for example, should participants in an experiment accept the experimenter’s claims about the location of responsibility as authoritative? Scientific experimenters have no special expertise in determining where moral responsibility lies.)

Suppose again that Jim, using the budget model scanner, has determined that the Captain will now release nineteen Indians if someone else kills one, that there is a high probability but no certainty that, if approached, the Captain will offer to free the remainder if one is killed, but that he is also experiencing rising irritation and so may not make an offer until he becomes willing to spare only eighteen, or even fewer. If whether Jim will be responsible for the killing depends on whether the Captain makes an offer, and if whether or to what extent Jim would be responsible is relevant to the permissibility of his engaging in killing, then the probabilities could be such that on Kamm’s view Jim

ought to approach the Captain rather than kill one Indian now, even though if he does that he may have to kill two or more Indians rather than only one. But what *reason* might there be for him to take this risk? Whether he acts on his own initiative on the evidence of the scan or whether he acts on an offer from the Captain, he will act in exactly the same way for exactly the same reason. An offer from the Captain would not give Jim any additional reason to kill an Indian, would not strengthen any reason he already has, and would not rebut any moral objection to the killing of an Indian. Whether there is an offer from the Captain seems irrelevant to his deliberations. One might indeed make a point here similar to one that is often urged against the relevance of an agent's intentions to the permissibility of her action – a point to which I think Kamm is sympathetic – namely, that where responsibility for the consequences of an act will lie is not among the reasons that count either against or in favor of the act. Who will bear the responsibility may be *determined* by considerations that are relevant to the agent's deliberations, but is not itself an additional consideration to be taken into account in assessing the permissibility of the contemplated act.

Another way to think about the permissibility of Jim's killing an Indian is to imagine that Jim knows all the facts except what the source of the threat to the twenty Indians is. He knows that twenty Indians will be killed *somehow* unless he kills one of them, in which case the other nineteen will survive. Suppose Jim reasons, correctly in my view, that it would be permissible to intervene if the threat to the twenty were a natural event. He might then consider that although many people believe that it is more important to prevent harms caused by wrongdoing than to prevent the same harms from natural causes, no one supposes that it is more important to prevent naturally caused

harms than to prevent equivalent harms caused by wrongdoing. He could then conclude that it is permissible to intervene whether the threat is of natural or human origin – that is, that it would be permissible to kill an Indian even if the threat were from someone like the Captain. But permissible is permissible: the killing could not become “more permissible” with the addition of an offer.

I mentioned earlier another case in which it may seem intuitively more plausible to suppose that there is a transfer of responsibility from Agent to principal that affects the permissibility of the Agent’s action: namely, the case of the landlord and the lawyer. It is an assumption of the example that it is morally wrong for the landlord to evict the impoverished tenant, yet no one supposes that the lawyer is blamable or liable (to defensive action, punishment, or payment of compensation) for evicting the tenant in his capacity as the landlord’s Agent. I am, however, skeptical of Kamm’s treatment of this case as well. If the lawyer evicts the tenant, he may not be responsible for the harm caused *in the sense* that he may be neither blamable nor liable. But contrary to what Kamm says, it does not follow from this that the lawyer acts permissibly in carrying out the eviction. *If* it is true that he is neither blamable nor liable, that is not because his acting as an Agent or substitute actor transfers all responsibility to the landlord. It is, rather, because even though he acts impermissibly, he is nevertheless excused by the requirements of his professional role, a role that it was permissible for him to undertake.

Kamm rightly says that the landlord has a “right to do what is wrong.” But a right to do wrong is not a liberty-right or permission to do wrong. That would be a contradiction: a permission to do what is impermissible. The landlord’s right is only a claim-right, a right not to be prevented from evicting the tenant. And that right does

transfer from the landlord to the lawyer. But because the landlord lacks a moral permission to evict the tenant, there is no permission he can transfer to the lawyer to render the lawyer's act permissible. Moreover, the eviction has no justification of the sort that Jim's killing an Indian has. The only positive reason for the lawyer to evict the tenant derives from his contractual duty to the landlord. Yet this duty may not be sufficiently strong to override the reason not to evict the tenant; nor is it binding, for the lawyer can resign. If he has been paid in advance for services including the eviction, he should refund the payment. The landlord would then have no justified complaint, for he cannot regard the lawyer's prior commitment to work for him as binding when what he demands is impermissible.

It seems to me, therefore, that neither of Kamm's central claims in her chapter on responsibility is correct. People who act as substitute actors or delegated agents of others have greater responsibility for the bad effects of their vicarious action than Kamm supposes. The two claims I have challenged are, however, only the minutest fraction of the rich body of argument in this magnificent book.

ⁱ Frances Kamm, *Intricate Ethics* (New York: Oxford University Press, 2006), p. 311.

All subsequent page references to this book will be in parentheses in the text.

ⁱⁱ Stanley Milgram, *Obedience to Authority* (New York: HarperCollins, 1974), p. 160.

ⁱⁱⁱ Alan M. Dershowitz, *Reasonable Doubts* (New York: Simon and Schuster, 1996), p. 177.