

# A PROBLEM FOR THE DOCTRINE OF DOUBLE EFFECT

by Sophia Reibetanz

**ABSTRACT** The Doctrine of Double Effect has been defended not only as a test of character but also as a criterion of wrongness for action. This paper criticises one attempt to justify the doctrine in the latter capacity. The justification, first proposed by Warren Quinn, traces the wrongness of intending harm as a means to the objectionable features of certain reasons for making this our intention. As I argue, however, some of the actions which seem to us to be permissible, and whose permissibility the DDE is supposed to explain, can be performed for these objectionable reasons. Since the proposed justification implies that any action is wrong when performed for these reasons, it renders the DDE incapable of accommodating the very intuitions about action which its proponents would have it explain.

**M**uch recent discussion of the Doctrine of Double Effect (DDE) has been concerned with its adequacy as a criterion of wrongness for action.<sup>1</sup> The DDE was not originally formulated to serve this purpose: its roots lie in the appraisal of character, not the appraisal of action. But, when applied to action, it can seem to capture some widespread beliefs about differences in the moral status of actions which result in the same good outcome. Consider, for instance, the Trolley and Transplant Cases discussed by Philippa Foot and Judith Thomson.<sup>2</sup> In the Trolley Case, we can save five people from an oncoming trolley only by diverting the trolley onto a branch line where it will surely kill someone else. In the Transplant Case, we can also save five people—but only by removing a sixth person's vital organs and transplanting them into the five. Most of us believe it would be wrong to save the five in the Transplant Case, but permissible to do so in the Trolley Case, even though this would result in no less harm to the sixth person. And it has seemed to some that the DDE captures this conviction.<sup>3</sup> According to the DDE, it is wrong to harm some in the course of benefitting others if the harm is intended as a means to production of the benefit, but not wrong if the same harm is simply foreseen as a side-effect of this production (and the benefit is duly proportionate to the

1. See, among others, Kagan 1989, pp. 128–82; Kamm 1992, pp. 376–81; Quinn 1993; and Bennett 1995, pp. 916–200.

2. These cases are elaborated in Foot 1978, pp. 23–24 and Thomson 1990, pp. 135 and 176.

3. See Foot 1985, pp. 25–26. The view that the DDE can at least accommodate our intuitions about these Trolley and Transplant Cases seems to be shared even by critics of the doctrine, some of whom point to counterintuitive results only in other cases. See, for instance, Kagan 1989, pp. 132–65.

harm).<sup>4</sup> If we saved the five patients in the Transplant Case, we would have to intend the removal of the other person's organs as a means of saving the five. But, since the trolley's collision with the person on the branch line is not itself necessary to save the five others, we could simply foresee this harm as a side-effect of saving them. So the DDE implies that it is wrong to save five lives in the Transplant Case, but not wrong to do so in the Trolley Case. It appears, then, to accommodate our beliefs about the different moral status of these actions. And in drawing our attention to a difference in the agent's intentions, it seems to locate a plausible source for this difference in status.

If the justification of moral principles were merely a matter of intuitive fit, then the adequacy of the DDE could be assessed solely by adducing such considerations as these. Both critics and proponents of the doctrine, however, have thought it necessary to search for some independent rationale for the doctrine. This paper is concerned with one such rationale. The rationale can seem to be one of the more successful of the attempts to justify the DDE. But it is my contention that it in fact renders the doctrine incapable of accommodating the very intuitions about right and wrong action which it is supposed to support—intuitions like those which we have about the Trolley and Transplant Cases. The rationale traces the wrongness of intending harm as a means to the objectionable features of certain reasons for doing so. As I shall argue, however, even certain actions which seem to us to be permissible, such as diverting the trolley and deciding not to remove the one person's organs, can be performed for these objectionable reasons. An appeal to features of our reasons for action may successfully rationalise the DDE as a test of character. But it cannot, I claim, successfully support it as a criterion of wrong action.

## I

When asked to give a rationale for the DDE, some have responded simply: 'To intend to harm someone as a means is to intend to *use* that person. And the relation of user to used is not one which it is morally appropriate for us to take up toward other persons. This is why it is wrong to intend harm as a means, but not wrong simply to foresee it as a side-effect of our action.'<sup>5</sup> In its present form, this proposal is inadequate. To use someone just *is* to treat him as a means; so we cannot show why it is wrong to harm someone as a means by appealing to the wrongness of using others. Furthermore, the notion of treating others as a means is importantly broader than the notion of harming others as a means—one can treat others as a means by *benefitting* them as a means.<sup>6</sup> So not only does the proposed

4. By 'harm' I mean 'substantive harm', not mere 'lack of benefit'.

5. Such a response is given by Fried 1978, pp. 28–29.

6. As is pointed out by Kagan 1989, pp. 178–82.

justification fail to do more than re-describe the kind of treatment which the DDE deems objectionable; it re-describes this in such a way as to imply, implausibly, that it is also wrong to benefit others as a means.

Nevertheless, the proposed rationale seems to point in the right direction. What we need, it seems, is a more precise statement of the way in which the relation of user to used enters into the thought of the agent who intends to *harm* people as a means. One such statement has been offered by Warren Quinn. Quinn claims that, when an agent intends to harm others as a means, he exhibits 'a distinctive attitude toward his victims'—namely, that they exist 'then and there *for his purposes*'.<sup>7</sup> Such an agent sees those whom he harms simply as tools for the achievement of his purposes, and not as beings with purposes of their own, purposes equally deserving of respect. Let us call this attitude the *user attitude*.

It is clear—though Quinn himself does not make this explicit—that we can intend to benefit others as a means without taking up the user attitude. For those whom we benefit as a means are normally affected by us in ways which aid in the achievement of *their* purposes. Consequently, even though we benefit them only as a means to the achievement of our own purposes, we need not thereby regard them as if they existed only for our own purposes.

We seem, then, to have arrived at a rationale for the DDE which circumvents my earlier objection while preserving the intuitive force of the original suggestion. We can state the rationale as follows: 'It is wrong to act on any intention to the extent that this intention is sustained and explained by the user attitude; for this is a morally objectionable attitude. Since one would have reason to intend harm as a means only if one held the user attitude, it is always wrong to intend harm as a means.'<sup>8</sup> The rationale, then, appeals to the moral inappropriateness of the user attitude, and claims that because this attitude is a morally inappropriate one, so too is any intention which, like the intention to harm as a means, a given agent would not have reason to adopt but for the user attitude.<sup>9</sup>

## II

Consider now the following elaboration upon the Transplant Case.

Rather than harming one person as a means of saving five others through transplants, the surgeon decides to let the five die. Some days later, a utilitarian friend asks why he responded in this way.

7. Quinn 1993, p. 190. Quinn is actually discussing a variant of the DDE which forbids what he calls 'direct harmful agency'—roughly, agency which enlists others in the realisation of the agent's purposes with the result that these others are harmed. But the difference between such agency and harming as a means is not important for our purposes.

8. Both here and in what follows, I use 'reason' to mean 'motivating reason'. I do not mean to imply anything about the normative reasons which agents might have.

Blushing, he replies, 'Had I been alone, I'd have had little compunction about removing the one's organs to save the five. But I was with a senior colleague who is a staunch defender of the DDE. I thought I'd stand a better chance at promotion if she didn't think I had acted wrongly.'

Most of us would maintain that although these motives do not reflect entirely favourably on the surgeon's character, the fact that he had them is irrelevant to the permissibility of his decision not to remove the one patient's organs to save the five. But can we maintain this if we endorse the DDE in the form that is supported by the proposed rationale?

Our surgeon did not intend to harm anyone as a means. But, as he admitted, he intended not to intend harm as a means solely in order that he might not offend the moral sensibility of his colleague. And this second-order intention seems to stand in the very relation to the user attitude which the proposed rationale for the DDE deems objectionable. It is sustained by, and in part explained by, the possession of the user attitude. For our surgeon would not have intended not to intend harm as a means *solely* as a means of not offending his colleague had he not held the user attitude. That is, he would not have had reason to adopt this second-order intention had he not had the user attitude. It seems, then, that if we are to reason consistently from the proposed rationale for the DDE, we must extend the DDE to embrace the following condition:

- (i) the agent must not intend solely as a means not to intend harm as a means.<sup>10</sup>

9. That this rationale is very close to the rationale which Quinn had in mind for his variant of the DDE is suggested by his claim that 'what is specifically amiss in relations of direct harmful agency is the particular way in which victims enter into the agent's strategic *thinking*' and by his subsequent re-affirmation that the moral difference between direct and indirect harmful agency lies in the former's '*presumption that the victim may be cast in some role which serves the agent's goal*' (p. 190; my italics). But when Quinn later turns to discuss particular cases, he appeals not to the inappropriateness of merely presuming or thinking that victims may be treated in this way, but to the inappropriateness of so *treating* them. In light of these later appeals, we might read Quinn's earlier references to the 'additional presumption' and the 'distinctive attitude' not as references to a certain mental state which it is objectionable to act from, but rather as references to a certain way of acting which is simply objectionable in itself. If we read Quinn in this way, we arrive at a different rationale for the DDE. This rationale appeals not to the inappropriateness of acting from a certain attitude, but rather to the inappropriateness of treating others in a certain way—namely, involving them in the realisation of your own purposes without giving them the opportunity to consent. This rationale is not open to the objections which I shall raise against the first rationale. It suffers, however, from one of the same flaws as we saw earlier in the mere appeal to the inappropriateness of using others. For it would commit us, implausibly, to the wrongness of *benefiting* others as a means without giving them the opportunity to consent.

10. It might be objected that the idea of 'intending to intend' is incoherent. But we surely can, and sometimes do, reflect upon, and form resolutions concerning, our own intentions; and this is all that is presupposed by my talk of intending to intend. To intend to intend is simply to resolve to make something one's intention. (Note that I am not claiming that we must engage in such higher-order intentional activity in order to have a first-order intention; my suggestion is only that it is possible for us to do so).

Suppose now that:

The surgeon's friend points this out to him, explaining further, that it is this extended version of the DDE which his senior colleague endorses. The surgeon realises that since he is unable to mask his intentions, his senior colleague must have known that he did not adhere to the DDE in its extended form. He therefore resolves to lose his desire not to offend his senior colleague. And he resolves, further, not to intend harm as a means. Chuckling to himself, he thinks, 'Next time I'll be sure not to offend my colleague!'

If and when there is a next time, our surgeon will indeed satisfy the extended version of the DDE: he will not intend harm as a means and, since he has lost his desire not to offend his colleague, neither will he intend not to intend harm as a means solely as a means of not offending his colleague. However, his intending not to intend solely as a means not to intend harm as a means will have been undertaken solely as a means of not offending his colleague. And it will be true that the surgeon would not have had reason to adopt this third-order intention unless he had the user attitude. On pain of inconsistency, then, we must extend the DDE still farther to embrace the condition that

- (ii) the agent must not intend solely as a means (not to intend solely as a means not to intend harm as a means).

And what justification could be given for excluding from the doctrine conditions ruling out each intention of a still higher order which stands in the appropriate relation to the user attitude?

The difficulty here is not that we could not check whether this infinite regress of conditions is satisfied in any given case. Most of us simply do not have intentions of an order higher than the third order, so it seems likely that all of the conditions beyond (i) and (ii) will always be satisfied. The difficulty is rather that the very rationale for the means principle, the principle which appeared to render the DDE capable of supporting our reaction to the Transplant Case, appears now to support conditions which render the doctrine *incapable* of accommodating this reaction. For once the DDE incorporates (i) and (ii), it implies that our surgeon acts *wrongly* in refraining from harming the one.

And it is not only our reactions to the Transplant Case which the DDE fails to support. Suppose that:

On the day that our surgeon is first confronted with the choice of letting the five die or removing the one's organs and performing life-saving transplants, he walks dejectedly home from work, thinking: 'Well, I may have done something for my promotion, but I let five of my patients die when I could have saved them.' Presently, his route brings him alongside a trolley-line. He can just make out five

tiny figures moving in the distance, and one figure moving at the far left of these five, where the main line has branched into a subsidiary line. Suddenly, he hears a rumbling noise and cries of 'Stop that trolley!' The surgeon would not ordinarily trouble over such shouts; these trolleys often malfunction, and there is usually someone else able to help. But today, swept up by a surge of anger at himself for not having saved his five patients, he rushes to the switch by the trolley-track and throws it, diverting the runaway trolley onto the subsidiary track.

In diverting the trolley, our surgeon intended to save the five. But he would not have had this intention had he not felt a need symbolically to dissociate himself from his earlier response of refraining from saving the five patients. And he would not have had this need to dissociate himself from his earlier action if he had thought that the one healthy person could not legitimately be used as a means of saving the five patients. So the user attitude sustains, and in part explains, his intention to save the five from the trolley. How, then, can the proponent of the DDE claim that the surgeon acted permissibly?

Since the proposed rationale for the DDE traces the wrongness of intending harm as a means to the character of the attitude that sustains and helps to explain this intention, it seems to commit us to the view that *any* action is wrong when performed with an intention that is sustained by, and in part explained by, this attitude. As my examples show, however, even the actions which we think permissible can be performed with intentions that stand in the relevant relations to the user attitude. So the proposed rationale renders the DDE incapable of accommodating the very beliefs about action which it was supposed to capture.

### III

It appeared, earlier in our discussion, that we could escape the inadequacies of a rationale which appealed only to the moral inappropriateness of using someone by appealing instead to the objectionable features of the attitude that gives one reason to intend harm to others as a means. As I have tried to show, however, this focus on our reasons for adopting certain intentions and upon the attitude that underlies them renders the doctrine incapable of supporting many of our beliefs about right and wrong action. It may indeed be an appropriate focus for the rationale of the DDE as a test of character. For what seems most relevant to questions about the moral acceptability of someone's character is indeed not the mere content of her intentions, but her reasons for making them her intentions. Where this focus leads the DDE into difficulties is in the appraisal of action. For, once we trace the wrongness of actions to the agent's actual reasons for intending to perform them, our DDE can

leave no room for the idea of an agent's performing the right action for the wrong reasons. It must claim that any act is impermissible when performed with an intention which the agent would not have reason to adopt were it not for his possessing the user attitude. And when it makes this claim, it becomes unable to accommodate some of the very intuitions about action which its proponents would have it explain.<sup>11</sup>

*Department of Philosophy  
Emerson Hall  
Harvard University  
Cambridge, MA 02138*

#### REFERENCES

- Bennett, Jonathan (1995), *The Act Itself* (Oxford: Clarendon Press).  
 Foot, Philippa (1985), 'Morality, Action and Outcome', in *Morality and Objectivity*, ed. Ted Honderich (London: Routledge) pp. 23–38.  
 Foot, Philippa (1978), 'The Problem of Abortion and the Doctrine of the Double Effect', in *Virtues and Vices* (Berkeley: University of California Press, 1978) pp. 19–32.  
 Fried, Charles (1978), *Right and Wrong* (Cambridge, Mass.: Harvard University Press).  
 Kagan, Shelly (1989), *The Limits of Morality* (Oxford: Clarendon Press).  
 Kamm, Frances (1992), 'Non-consequentialism, the Person as an End-in-Itself, and the Significance of Status', *Philosophy and Public Affairs* 21: 354–89.  
 Quinn, Warren (1993), 'Actions, Intentions, and Consequences: The Doctrine of Double Effect', in *Morality and Action* (Cambridge: Cambridge University Press) pp. 175–93.  
 Thomson, Judith Jarvis (1990), *The Realm of Rights* (Cambridge, Mass.: Harvard University Press).

11. I am very grateful to Robert Audi, Brad Hooker, Derek Parfit, Tim Scanlon, and Helen Steward for criticisms which have helped to shape the paper.

