

The Doctrine of Double Effect: Intention and Permissibility

William J. FitzPatrick*

Department of Philosophy

Abstract

The Doctrine of Double Effect (DDE) is an influential non-consequentialist principle positing a role for intention in affecting the moral permissibility of some actions. In particular, the DDE focuses on the intend/foresee distinction, the core claim being that it is sometimes permissible to bring about as a foreseen but unintended side-effect of one's action some harm it would have been impermissible to aim at as a means or as an end, all else being equal. This article explores the meaning and application of the DDE along with current debates over the nature of the intend/foresee distinction and its moral significance. How is the line between intended effects and merely foreseen but unintended side-effects to be drawn, and how are problem cases best handled? What is an appropriate methodology for debating the tenability of the DDE? How might the DDE interact with other non-consequentialist principles, and how might it be modified to capture other related factors (resulting perhaps in a Doctrine of Triple Effect) or to avoid certain problem cases? Does the DDE make permissibility turn on the actual intentions of particular agents, as critics such as Thomson and Scanlon assume, or is the role of intention more abstract?

1. Introduction

The Doctrine of Double Effect (DDE) is a claim about the moral significance of the distinction between *intending* harm (as a means or as an end) and bringing about harm as a *foreseen but unintended side effect* of one's action. Specifically, the core claim of the DDE as it is currently understood is that *it is sometimes permissible* to bring about as a foreseen but unintended side-effect of one's action some harm it would have been impermissible to aim at as a means or as an end, all else being equal.¹

Imagine a pair of cases equalized for factors such as motive, consequences, causal structure (e.g., whether the harm is done or merely allowed), and so on, such that they differ *only* in that one involves intended harm while the other involves merely foreseen harm. For example, in Terror Bombing, a pilot carries out a targeted strike on a densely populated village as a means to sow terror and demoralize the enemy, hastening a favorable end to the war; by contrast, in Tactical Bombing, the pilot carries out a tactical strike on a munitions plant *despite* foreseeable harm to the same number of neighboring civilians (which he takes all reasonable measures to minimize), because of the military advantage of eliminating the munitions plant and similarly hastening a favorable end to the war. The DDE implies that at least sometimes, in pairs of cases like this, the latter action can be permissible even while the former remains impermissible, the difference stemming from the difference in intention: the harm to civilians is intended as a means in Terror Bombing (it is part of the bomber's plan for achieving his end) but is merely a foreseen but unintended side-effect in Tactical Bombing.

Proponents of the DDE differ in how strongly they understand the presumed moral prohibition against intending the death or serious harm of innocents as a means or end. Some, particularly within traditions of moral theology, take it to be absolute, allowing no exceptions (Anscombe 1981). Others take it to be defeasible by certain other factors, possibly including consequential considerations above some high threshold, such that in rare, extreme circumstances even an act such as Terror Bombing might be permissible to avoid truly catastrophic harm. The point, however, is that all proponents of the DDE agree that there will be many cases (below any such threshold) where it remains impermissible to bring about certain harms as an intended means or end while it is nonetheless permissible to bring about similar harms as an unintended side-effect. What is key is this possibility of an asymmetry in permissibility that turns on intention.

Now there are several points to notice straightaway. First, the DDE does not imply that it is generally permissible to bring about a harm in the course of pursuing a good end as long as the harm is a merely foreseen but unintended side-effect of one's action. Such a claim would be a non-starter, as many such cases would involve violating people's rights and/or bringing about harms that plainly outweigh the good achieved.² The DDE says only that *some* cases of bringing about foreseeable harm are permissible even where it would remain impermissible to aim at such harm as a means or end.

Related to this point, but often overlooked, is the fact that the DDE does not imply that in *every* pair of *otherwise similar cases* the intend/foresee distinction will make a difference to permissibility; instead, it allows for context-sensitivity in the moral force of the distinction. There may well be contexts—involving certain combinations of normatively relevant factors—in which the distinction makes no moral difference at all, *even* where all else is equal. For example, some defenders of the DDE hold that in circumstances in which a patient's life has irremediably ceased to be a good to her due to unmanageable and overwhelming suffering, and she has requested euthanasia, the usual moral worries about intending the death of an innocent person do not apply, so that the distinction between intending and merely foreseeing death may fail to make any moral difference in this context.³

The moral significance of the intend/foresee distinction may thus be *context-sensitive* in the same way claimed by most defenders of a distinct but parallel doctrine concerning the doing/allowing distinction, sometimes called the 'Doctrine of Doing and Allowing' (DDA). According to the DDA, it is *sometimes* permissible to *allow* a certain harm it would have been impermissible to *do*. This again is understood by its defenders to imply only that in *some* otherwise similar pairs of cases the doing/allowing distinction will make a moral difference; it does not imply that in all such pairs it will make a difference, and so it is compatible with the claim that in some contexts it will not, even where all else is equal between the cases (Foot 1994, Quinn 1993a, Kamm 1996, 2007.⁴).

This brings out an important *methodological* point for debates over both the DDE and the DDA that was largely overlooked until very recently: the DDE and DDA cannot be refuted by citing one or more pairs of cases where everything else is equal and yet the distinction in question fails to make any difference to permissibility. Contemporary defenders of the DDE and DDA typically *grant* the possibility of such cases from the beginning, for the reasons just noted. So while they do have a burden to explain *why* their distinction makes a difference in some contexts and not in others, they needn't be disturbed by the fact that in some circumstances, with certain combinations of normatively relevant factors, it doesn't.⁵

Another point to emphasize is that since the DDE, like the DDA, implies that the permissibility of a pair of acts can differ even while the consequences are held constant, the

DDE is plainly a *non-consequentialist* principle. It says that something other than the consequences of an act can matter in determining its moral status as permissible or impermissible: specifically, whether a harm is *intended* or *merely foreseen* can make a crucial difference to the act's moral status. Many non-consequentialists embrace both the DDA and DDE, taking them to complement each other in accounting for various non-consequentialist intuitions.⁶ The DDE by itself, for example, tells us nothing about what is wrong with using a machine to manufacture a drug to save five when as a foreseen side-effect it pumps lethal fumes into a room containing one who cannot be moved (Foot, 1978); the DDA, however, can help to explain this in terms of its typical prohibition against doing harm even for good ends (specifically, initiating a harmful sequence against innocents, which tends to violate negative rights). At the same time, however, the DDA fails to discriminate between Terror Bombing and Tactical Bombing, since *both* involve doing rather than allowing harm; it is precisely for cases such as this that the DDE is thought by its defenders to be crucial as part of a more comprehensive non-consequentialist view.

What the DDE suggests here is that even if there are exceptions (e.g., in wartime) to the DDA's prohibition against initiating a lethal sequence against innocents, *those exceptions will be limited* by the DDE's prohibition against intending the deaths of innocents as means. Our two contrasting cases bear this out: while it may, under certain circumstances, be permissible to perform an action (such as bombing a munitions plant) despite the fact that this initiates a lethal sequence against innocents who suffer the fallout, it remains crucial according to the DDE that this be 'collateral' damage—a foreseen but unintended side-effect of the action, rather than an intended means toward one's end, which latter remains impermissible despite the advantages that might be had through targeting civilians. The DDE is thus often seen as important in navigating between pacifism, on the one hand, and terrorism on the other, allowing a principled way to reject the first without thereby sliding into justification for the second (McMahan 2009, Anscombe 1981). As we will see, however, not all non-consequentialists take this view, and some reject the DDE even while accepting the DDA or other non-consequentialist principles, finding other ways to account for the intuitions the DDE has been used to explain (Thomson 1985, Kamm 2007, Scanlon, 2008).

In the remainder of this article we will consider some recent controversies and developments, starting with questions about the intelligibility and application of the very distinction at the heart of the DDE. This is not meant to be comprehensive, but simply to explore some of the main issues currently being debated.

2. The Intend/Foresee Distinction

Many have attacked the DDE on the grounds that the distinction between intended effects and merely foreseen but unintended side-effects is so vague and arbitrary that virtually any case of what intuitively seems to be intended harm can be recast as a case of merely foreseen, unintended harm, rendering the distinction morally useless: if it can be used to justify *anything* then it can be so abused to justify virtually *everything*. Consider the Terror Bomber again. What stops him from appealing to the double effect distinction himself as he drops his bombs on a major population center (Bennett 1981, 1995)?⁷ He might, after all, say:

True, my intended means to ending the war is to demoralize the enemy and bring about their capitulation, but for that I don't actually *need* the civilians to be *killed*: I just need them to *appear* dead for a while, which is all I'm intending as my means; it would be fine with me if my

bombs merely knocked them out for a couple days from the impact (the shrapnel and debris miraculously missing them all), until their government's surrender. Of course, I know this is extremely unlikely, and so I *foresee* their likely deaths, but this is just a regrettable, unintended side-effect of making them appear dead.

In another (hypothetical) example, a group of cave explorers are trapped when their largest member attempts to exit first and gets hopelessly stuck in the exit even as waters are rising inside (Foot, 1978). Suppose the group has dynamite and uses it to blow the man to bits, thus clearing the passage and securing escape, and then they attempt to justify themselves as follows:

Our end was simply to escape, and our means was to clear the opening. Granted, the means by which we cleared the opening was to blow him to bits, but we didn't intend his *harm* as such. That was just a foreseen but unintended side-effect of blowing him to bits.

Again, the objection is that if we can claim not to have intended harm *even* in a case like this, then the intend/foresee distinction is too easily manipulable to do real moral work.

The two cases above actually raise different issues. It turns out that the first is easily dealt with, as it just embodies a confusion about the structure of the case. The reply to the bomber is that what he's taking to be a side-effect is in fact just *a more proximate means*: while we can grant that his means to demoralize the enemy is to make the citizens appear dead to the government, the obvious question is *by what more proximate means* he aims to bring *that* about. Since he does this by dropping lethal bombs on them, with no means of directing the bombs merely to knock them out, the answer is: he brings about their appearing dead *by killing them*, which is a very effective way—and the only way open to him, given the tools he is using—of making people appear dead. It's beside the point that he would be perfectly happy if by some miracle no one actually died and all were just knocked out for a while: all that shows is that their deaths were not aimed at *as an end*, and that there were other conceivable ways in which his end could have been achieved without their deaths. None of this shows that he wasn't *in fact* aiming at their deaths as *his way of making them appear dead*. Plainly he was. Thus, it's just false that their deaths were mere side-effects: they were simply the more proximate means by which he made them appear dead. Defenders of the DDE therefore needn't worry about such sophisticated abuses of the intend/foresee distinction.

The second case has a different structure and raises a more difficult problem. Here there is no conflation of more proximate means with side-effects: the cave explorers do not blow up the man *by* killing him, so killing him is not their intended means of blowing him up. Still, we may wish to say that they nonetheless intend his death as a means to escaping inasmuch as they so intend his being blown to bits! After all, denying this might well give the impression of reducing the DDE to a sophisticated game in which we are free to 'direct our intention' as we wish, focusing on whichever description of our action we prefer ('making civilians appear dead', 'blowing the man to bits') and calling that 'our intention', while arbitrarily denying that we intend what we are doing under less palatable descriptions ('killing civilians', 'killing the man blocking the exit'). This would indeed make nonsense of the DDE from the very start (Anscombe 1981, 59; 2001).

Some defenders of the DDE have nonetheless been led to embrace the cave explorers' highly counterintuitive claim above, and others like it, though on more principled grounds, such as a narrow or "strict" construal of intention. Lawrence Masek (2010), for example, holds that anything lying outside the intended end and the narrowly construed causal sequence leading to it is strictly unintended (569). Thus, along with accepting the

cave explorers' claim, he holds that one could legitimately claim not to be intending the death of a person one forcibly vivisects to harvest organs for use in saving others (though he does not claim that this suffices to justify the action): all that is intended as a means while carving the person up is the removal of organs to make them available for others (570). The challenge for such an approach (setting aside natural worries about the very intelligibility of such claims about intention), of course, is to show how the DDE will still be a useful principle if it's *that easy* to carve off harms as mere 'side-effects': the double effect distinction will turn out to apply to so many cases—far more than those in which it might plausibly be thought to do real justificatory work—that its usefulness for purposes of principled moral discrimination becomes unclear.⁸

The more standard move for defenders of the DDE is to reject such counterintuitive claims and to seek a way to block such extensions of the double effect distinction.⁹ We should not accept the cave explorers' claim, for example, because the relationship between the man's *being blown to bits* and his *being killed* is 'too close' for it to make sense for his colleagues to say that they aim at the former but not at the latter. But the challenge is to articulate what exactly this 'excessive closeness' amounts to, offering a principled account that rules out the cave explorer's claim but doesn't at the same time rule out plausible cases of double effect (Foot 1978, 21–2). Jonathan Bennett (1995) surveys a number of possible accounts in detail and finds them all unsatisfactory, concluding that the intend/foresee distinction simply cannot do the principled moral work it would have to do for the DDE to be plausible and useful.

Defenders of the DDE, however, have developed solutions to the 'problem of closeness' not considered by Bennett, which they believe deal satisfactorily with the problem. According to one such account, for example, the key is to distinguish between *constitutive* and merely *causal* relations between *states of affairs*: where the relation is a merely causal one, there is room in principle to speak intelligibly of aiming at one but not at the other; where the relation is constitutive, however, there is not, i.e., it is 'too close' to allow for double effect claims (FitzPatrick 2006). In the cave case, the relationship between *his being blown to bits* and *his being killed* is too close because it is a constitutive relationship—his being blown up just constitutes his being killed—so that there is no room for intending the first but not the second.¹⁰ By contrast, if a runaway trolley is headed toward five trapped people and you turn it onto a side track on which only one is trapped, you can intelligibly be said to aim at *the trolley's being diverted* toward lesser harm without thereby aiming at *the one's being killed* down the side track (though this is a foreseen side-effect of the diversion), because here the relation is a merely causal one (the trolley's being diverted away from the five causes the one's being killed down the side track, but is not constitutive of it).¹¹

If this or some other account of the intend/foresee distinction (see e.g., Cavanaugh 2006, chap. 3, Hills 2007) yields intuitively correct answers in a wide range of cases, it may provide a principled way of putting the brakes on overextensions of the double effect distinction, rescuing the DDE from the above worries about arbitrariness and sophistry.¹²

3. Quinn's Modified DDE

A different response to such worries about the double effect distinction has been simply to modify the DDE to side-step the need for an account of excessive closeness. This is the approach taken by Warren Quinn (1993b). Whereas the traditional DDE discriminates (in some contexts) against the intention of *harm* as a means or end (in favor of foreseen but unintended harm), Quinn shifts and broadens the focus: what is really

problematic is not just intending harm as such but intending someone's *involvement* in some event (without his consent), as a means to one of the agent's ends or as an end itself, where this involvement is known to the agent to be harmful to the person. According to Quinn, this intentional relation is already problematic even if 'the harm itself' is not strictly intended: for such an agent still takes a "strategic view...and distinctive attitude toward his victims...[treating them] as if they were then and there for his purposes"—as if they were "material to be strategically shaped or framed by his agency" (190).¹³

Thus, even if the Terror Bomber were granted his claim that he does not intend harm to his victims (he intends only to make them appear dead), Quinn claims that he still runs afoul of the modified DDE: for he clearly does intend the *involvement* of his victims in something as a means to his end, where that involvement is foreseen to be harmful to them, and that is enough to be problematic. Similarly, we needn't develop an account of closeness to see that the cave explorers' action is an instance of the kind of agency discriminated against by Quinn's modified DDE: even if they don't strictly intend 'the harm itself' to the man they blow up, they clearly intend his involvement in the explosion as a means of clearing the cave, foreseeing that this involvement will be harmful to him.

Still, while Quinn's modification of the DDE does provide a way around certain problem cases, it's not obvious that such a departure from the traditional focus on intention of harm is necessary, since as we have seen there are other ways of handling the alleged problem cases. Some defenders of the traditional DDE will also object that the shift in focus loses something important about intending *harm or evil as such* (Kamm notes that Nagel 1986 would likely have this reaction, given his rationale for his view), and critics have raised further worries about whether the modified account really yields plausible verdicts in a variety of cases (Bennett 1995, Cavanaugh 2006, 59–66, FitzPatrick 2006, Kamm 2007, 82–9).

4. *The Loop Trolley Case: Is Intention of Harm Really the Issue?*

Setting aside difficulties with the notion of intending vs. merely foreseeing harm, many have questioned whether this distinction is really morally significant in any case—at least in the way claimed by the DDE, i.e., as sometimes making a difference to permissibility. Consequentialists, of course, reject that claim of moral significance (though they might in principle hold that an agent's intending of harm makes the overall state of affairs that includes it worse, thus allowing the double effect distinction a role in the theory of the good, if not in the theory of right). But as noted earlier, even some *non-consequentialists*, who accept constraints against certain actions despite their great utility, have denied that this ever has to do with the intention of harm as a means or end, as claimed by the DDE. One argument to this effect stems from a variation on Foot's (1978) famous Trolley Case: the Loop Case introduced by Judith Jarvis Thomson (1985).

As before, a runaway trolley is heading toward five ahead but you can divert it onto a side track containing only one.¹⁴ The difference is that now the track beyond the one curves back around to the main track in such a way that if the trolley were to continue on that path it would loop back and hit the five from behind. Diverting the trolley thus does no good by itself. But suppose the one on the looping side-track is so large that the trolley's hitting him will stop it, preventing it from looping around and hitting the five. Now diverting the trolley will again save the five. Thomson argues that it is permissible to turn the trolley in this case: how could it *not* be permissible here if it's permissible in the original trolley case, since this case differs objectively only in the presence of some

extra track beyond the one, which is *never even used* in any case (since the trolley stops when it hits the one, before even beginning to loop back)? But if this is permissible, then apparently the intention of harm as a means doesn't make for impermissibility here, since we seem to intend the death of the one as a means to save the five (since without the one's being there to be crushed there would be no point in turning the trolley). Moreover, if intention of harm does not suffice for impermissibility, as the Loop Case suggests, then that means that one needn't generally establish the *absence* of intention of harm in order to establish *permissibility*. And that in turn casts doubt on whether the absence of the intention of harm was really the crucial factor even in the original Trolley Case, as proponents of the DDE claim (since again, if the analysis of the Loop Case is correct, absence of intention of harm is not crucial, so that whatever else it is that makes turning in the Loop Case permissible may also be what really made turning in the Trolley Case permissible to begin with, rather than absence of intention of harm).

Many take the Loop Case to undermine the moral significance of the double effect distinction. Notice, however, that such a conclusion would be too hasty even if one accepted Thomson's conclusion about the permissibility of turning in the Loop case. All that follows is that *in this context* (of diversion of a public threat) intention of harm fails to be *sufficiently* morally problematic to make for impermissibility. This is compatible both with (i) the claim that even here these factors contribute *some* moral force, making Loop at least more problematic than Trolley (Wedgwood, 2011), and with (ii) the claim that in *other* contexts, with different combinations of factors present, the double effect distinction *does* make a sufficient difference concerning permissibility (the principle of contextual interaction again). Showing that within the family of diversion cases intention of harm is not sufficient for impermissibility does not refute the DDE.

Defenders of the DDE might in any case reject Thomson's analysis of the Loop case. One might hold, for example, that her intuitive verdict about the case—that it is permissible to turn the trolley—is simply mistaken (Otsuka 2008, McMahan 2009). While it is true that the only objective difference between Loop and Trolley is the presence of some extra track, this *does* plausibly matter insofar as it changes what you're doing when you divert the trolley, and therefore affects your attitude toward the one: *diversion by itself no longer removes the threat from the five* as it did in Trolley (with the death of the one being a mere side-effect one wishes wouldn't happen), but instead just changes the direction from which the five are hit; it helps *only due to the one's being crushed*, which means that this has become a crucial means to saving the five, bringing us back to plausible worries about intention. (Imagine if it were instead a runaway car and one had to chase down the one—steering after him as he tries to escape—and crush him so that it would stop before hitting more. Is Loop really different in kind, rather than just degree?).

Another possible response is to grant Thomson's conclusion that the Loop Case is permissible while rejecting her claim that it strictly involves intention of harm. This brings us to another proposed modification of the DDE.

5. The Doctrine of Triple Effect

Frances Kamm (2007) has argued that in addition to (i) effects intended as a means or end, and (ii) foreseen but unintended side-effects, we need to recognize a third category: (iii) effects *because of* which we act, in the sense that we wouldn't act unless we expected them to occur, though we don't act *in order to* bring them about as in (i). It's easy to conflate (iii) with (i), but they're distinct and this can sometimes make a moral difference. Michael Otsuka (2008) helpfully illustrates the distinction between (iii) and (i) with

another variation on Trolley, the Six Behind One Case: as before, you can divert the trolley away from five onto a (non-looping) side-track with one person on it, but this time there are six *more* people behind the one on the side-track; fortunately, the one is large enough to stop the trolley before it reaches the six. Here you do not turn the trolley *in order to* crush the one: you turn it simply in order to save the five; but still, you turn it *only because* (i.e., on the condition that) the one is there, since otherwise you'd actually be doing more harm than good by turning (six deaths instead of five).

To take account of this factor and the moral difference it can make, Kamm thus proposes the 'Doctrine of Triple Effect' (DTE), which allows for some cases of harm in category iii (as well as ii, like the DDE) even where harm in category i is prohibited—as in the above example, where we may turn *because* the one is there but not *in order to* hit the one.¹⁵ This, she thinks, also helps to make sense of the permissibility of Thomson's Loop case without abandoning the DDE's ban on intending harm as a means: on her analysis, when we turn the trolley in the Loop case we do not actually do so *in order to* hit the one as a means, though it's true that we do so *only because* the one is there. Otsuka (2008) argues, however, that although Kamm's distinction is a good one and applies to the Six Behind One Case, it doesn't in fact apply to the Loop case, which remains one of straightforwardly intending the death of the one as a means. On this view, then, Kamm is mistaken to think that the DTE supports the permissibility of turning in the Loop Case: turning remains impermissible there for the reasons given by the ordinary DDE (at least unless there are special contextual factors that undermine the usual negative force of intending death as a means).

6. Permissibility and the Actual Intentions of Particular Agents

The final controversy we will consider, which traces back (at least) to objections by James Rachels (1994, originally published in 1981) and Judith Jarvis Thomson (1999) but has recently been revived and given new prominence by Thomas Scanlon (2008), goes to the heart of the interpretation of the DDE. Rachels, Thomson and Scanlon all assume that the DDE makes the permissibility of some actions turn on the *actual* intentions of *particular, concrete* agents, and they go on to argue that this has absurd implications.¹⁶

Their strategy is to consider acts of the type usually judged permissible, such as tactical bombing, and to imagine that act done with the intention *usually* associated with the *impermissible* type of act, such as terror bombing. Suppose that the pilot selected to fly the sort of tactical bombing mission usually favored by defenders of the DDE turns out to harbor bad intentions: perhaps he hates the civilians in the neighboring village and delights in their foreseeable deaths from the fallout of the factory's destruction, rather than seeing them as regrettable side-effects; or perhaps he doubts the significance of the munitions plant but accepts the mission because he knows it will kill civilians and thinks this worthwhile as a means of demoralizing the population and lessening their support for the war. He won't modify the mission and target them directly, but he accepts the mission precisely as a way legally to kill some civilians; he aims at and welcomes their deaths as a means or as an end, even though he's outwardly doing nothing differently from any other pilot carrying out the tactical bombing mission against the munitions plant.

According to the interpretation of the DDE assumed by Rachels, Thomson and Scanlon, the very bombing mission (targeting the munitions plant) that would have been permissible with a different pilot is rendered *impermissible* by this pilot's wrongful intentions—though it would be rendered permissible again if a different pilot with proper intentions took over. But this, the objection goes, is absurd: if it is permissible to fly a

tactical bombing mission against this military target in the overall circumstances having to do with its military value (in the broader context of the great value of ending the war quickly and favorably) and the proportionality of that value to the foreseeable collateral damage, then surely nothing that is going on inside any particular pilot's head can make that mission impermissible and tie our hands.¹⁷ As Rachels puts the point: 'A pure heart cannot make a wrong act right; neither can an impure heart make a right act wrong' (Rachels 1994: 142). Just as with motives, "the intention is not relevant to deciding whether the *act* is right or wrong, but instead it is relevant to assessing the character of the person who does the act, which is very different" (141).¹⁸ The pilot with bad intentions may be a bad person and may *act badly* in flying the mission against the munitions plant, since he's acting with bad intentions (and perhaps bad motives as well); but the *act* in question—attacking the munitions plant—remains permissible.¹⁹

One response to this objection is to accept the critics' interpretation of the DDE as making permissibility turn on the actual intentions of particular agents, and to try to show that the implications in such cases are not after all implausible. If what we're talking about is the permissibility of a 'thin' act-type such as 'bombing the munitions plant' (where the intention with which the act is done is not built into the description), then the critics' denial that the permissibility of the act depends on anyone's actual intentions seems very plausible. But if instead we consider the 'thick' act-type 'bombing the munitions plant with the intention of killing nearby civilians'—which is of course what the bad-intentioned tactical bomber described in the objection would be doing—then it's not so obvious that *this* is permissible. After all, if—as even the critics concede—bombing with such intentions is *acting badly*, and so is something to be *avoided*, then that seems to support the claim that it is impermissible to perform this act (i.e., an act of this thick act-type), and so is impermissible for the bad-intentioned bomber to fly the mission. So understood, the DDE's claims about such cases needn't seem absurd at all (Wedgwood 2011, and for a similar account, McMahan 2009).²⁰

A different response to the objection raised by Rachels, Thomson and Scanlon is to reject their interpretation of the DDE from the start. Although it is common and usually harmless to speak loosely of what particular agents intend when applying and explaining the DDE, it may be a mistake to think that this implies that the DDE makes permissibility turn on the actual intentions of particular agents. An alternative understanding of the DDE construes it as a *constraint on the justification of action*, having to do *not* with particular agents' actual intentions but with the intentions that *would* be involved in performing a certain type of action for the reasons provided by a proposed justification (FitzPatrick 2003, forthcoming). The idea is (roughly) that it is a condition of adequacy on any justification J of action A that a good agent could *act well* in performing A for the reasons given in J. But acting in a way that involves intending harm to innocents as a means or end is *acting badly* (something even many critics grant). So if J is to be adequate, it cannot be such that an agent who does A because of J would thereby be engaging in the intention of harm toward innocents as a means or as an end. There is therefore a constraint on any justification J of action A such that A cannot be justified (shown to be permissible) by showing it to be an efficient means to a good end in cases where A constitutes such a means precisely *by virtue of* being *harmful to innocents*.

This account yields the intuitively correct answers about the original Terror Bombing and Tactical Bombing. The proposed justification for Terror Bombing—that bombing the village will kill civilians, demoralize the population and hasten a favorable end to the war—clearly fails the above condition of adequacy: no one could perform the action in question for the reasons given in that justification without intending the deaths of

innocents and thereby acting badly. So the proposed justification fails to show Terror Bombing to be permissible, and unless there is some other justification for it that survives the condition of adequacy (which there presumably isn't), Terror Bombing will be impermissible. By contrast, the justification for Tactical Bombing—that bombing the munitions plant will weaken the enemy's military capacity and hasten a favorable end to the war—satisfies the adequacy condition: a good agent could perform the action in question for the reasons given in that justification *without* intending the deaths of innocents. So at least if the justification is otherwise sufficient (in terms of issues of last resort, proportionality, etc.), Tactical Bombing may well be permissible. We thus have an account of the intuitions about these cases that makes no appeal to the actual intentions of the particular agents involved, but relies instead on abstract considerations having to do with justification, intention and acting well.

Now what about the pilot flying the mission against the munitions plant, like the Tactical Bomber, who happens to harbor the bad intentions of a Terror Bomber? The critics complained that the DDE implies that the mission against the munitions plant becomes impermissible in this case due to this particular pilot's bad intentions, despite the fact that it would otherwise be permissible—which seems absurd. On the current proposal, however, the DDE implies no such thing. The actual intentions of this particular pilot are irrelevant to the permissibility of flying the mission against the munitions plant: what matters is just whether there is an otherwise sufficient justification *J* for flying such a mission against the munitions plant (i.e., the same as in the original Tactical Bombing case), such that a good agent *could* fly such a mission for the reasons given in *J* without thereby intending harm to innocents and acting badly. This of course yields the same verdict on permissibility as in the original Tactical Bombing case: the mission against the munitions plant is permissible, period, for the reasons given in the justification and because a good agent *could* act on those reasons without being involved in bad intentions and acting badly. If this particular pilot flies the mission for the wrong reasons, with bad intentions and perhaps bad motives, *he* will act badly. But the DDE, on the present construal, doesn't make the *permissibility* of the act turn on whether *this particular agent* will act well: what matters to the permissibility of the act is just whether there is an adequate justification for it (such that an otherwise good agent *could* act well by so acting because of this justification), just as the critics themselves insist (FitzPatrick forthcoming).²¹

It seems, then, that defenders of the DDE have ample resources to provide plausible responses to the sort of objection raised by Rachels, Thomson and Scanlon. Even if the objection can be answered, however, it remains of interest insofar as it forces defenders of the DDE to clarify their views about how and why exactly intentions can matter to permissibility—in particular, whether actual intentions of particular agents affect the permissibility of acts or whether the primary role of intention is more abstract, affecting what counts as a legitimate justification for an act.

Short Biography

William J. FitzPatrick is Associate Professor of Philosophy at the University of Rochester, and holds an AB in Philosophy from Princeton University and a Ph.D. in Philosophy from UCLA. Prior to joining the faculty at the University of Rochester he taught at Virginia Tech and Yale University. His research centers on a defense and development of a robust, non-naturalist ethical realism and externalist theory of practical reasons, along with critiques of neo-Kantian constructivism and of recent appeals to evolutionary theory to debunk ethical realism. He has also done work in defense of non-consequentialist

normative ethics, and against skepticism about moral responsibility, along with applied work on justice and climate change, the value of nature, and ethical issues surrounding stem cell research. His publications have appeared in such journals as *Ethics*, *Mind*, *Analysis*, *Philosophical Studies*, *Oxford Studies in Metaethics*, and *Oxford Studies in Normative Ethics*.

Notes

* Correspondence: Department of Philosophy, University of Rochester, Box 270078, Rochester, NY 14627-0078, USA. Email: william.fitzpatrick@rochester.edu.

¹ The expression “double effect” refers to these two distinct effects of some harmful actions: (i) the effect one *aims* at (or more precisely the set of effects comprising one’s end and the means one intends toward that end), and (ii) the *foreseen but unintended* harmful side-effect (to be regretted and minimized), as illustrated by Tactical Bombing just below. For the historical development of the DDE, see Cavanaugh (2006, chap. 1).

² Philippa Foot (1978) illustrates this point with the case of some “wicked merchants” who sell oil they know to be poisonous, rather than discarding it, so as not to lose the profits from the sales. The resulting harm is foreseen but not intended as a means or end (they do not need the people to be harmed—it would be just as well if no one ever actually used the oil, as their harm is incidental to the merchants’ plan, which is simply to get profits from selling the oil); still, the merchants’ action is obviously impermissible: the good achieved is not worth the foreseen harm, and the action clearly violates the rights of the victims.

³ Consider two cases of high-dose morphine injection differing only in that one is a case of euthanasia, where death is intended as a means to end the suffering, while the other (adhering to Catholic strictures against euthanasia) involves only using the morphine to control the immediate pain despite foreseeing (but not intending) that the high dose will hasten organ failure and death. On the above view, these cases may be *morally equivalent* because they’re *equally permissible*. This, of course, goes against Catholic moral doctrine, which is the historical source of the DDE: the Catholic position is that the DDE continues to apply here, ruling out euthanasia (because of the implied intention of death) while allowing similar acts that can be justified without aiming at death. But there is nothing stopping contemporary moral philosophers from embracing the DDE itself while departing from some traditional moral claims about its application. See Quinn (1993b, fn 17) for a defense of such a departure in “non-conflict cases”.

⁴ This point was actually noticed in passing at least as far back as 1967 by Foot (in the article reprinted in 1978), though it was made explicit only later in response to challenges to the DDA by Rachels (1975) and others. Kamm has provided perhaps the most extensive discussion of the more general issues surrounding this point in terms of her Principle of Contextual Interaction.

⁵ A classic example of an attack on the DDA by appealing to a single pair of cases where the doing/allowing distinction apparently fails to make a moral difference is Rachels (1975). His argument fails because a theory such as Foot’s (1994) can readily explain how the killing/letting die distinction can make a difference in some contexts where all else is equal, but not in others, such as the pair of cases Rachels presents. In Foot’s cases of Rescue I and Rescue II, for example, her claim about the greater stringency of negative rights over positive rights explains how it can be permissible in Rescue I to allow one to drown while racing to save five others from drowning (the positive right to aid on the part of the one is overridden by the positive rights of the five), even though it is impermissible in Rescue II to run over one in a jeep in order to get to the five in time (the *negative* right of the one not to be killed is *not* overridden by the positive rights to aid on the part of the five). At the same time, she can also explain why her theory does not imply any moral difference in Rachels’ examples of Smith and Jones, where Smith drowns his young cousin in the bathtub for money while Jones merely lets him drown after he falls: this pair of cases is structurally very different from Foot’s pair, since unlike in Rescue I, the positive right of the child in the Jones case is *not* overridden by other considerations (Jones’ desire for undue gain does not override a child’s positive right to be saved), which means that no asymmetry in permissibility arises between the Smith and Jones cases. Similar lessons may be drawn for the DDE: there may well be some pairs of cases that are morally equivalent because equally permissible (as already noted in connection with voluntary euthanasia) or equally impermissible (where egregious causing of foreseen harm as a side-effect is no better than intentional harm), but none of this tells against the existence of *some* pairs of cases where the intend/foresee distinction makes for an asymmetry in permissibility. This point will come up again later.

⁶ While the DDA and DDE are certainly compatible, it’s important to see that they are distinct. The two relevant distinctions cut across each other: (i) one can *do* harm that is *intended* (e.g., a murder), but (ii) one can equally *do* harm that is *merely foreseen* (e.g., Tactical Bombing); and (iii) one can *allow* harm that is *merely foreseen* (e.g., Rescue I) but (iv) one can also *allow* harm that is *intended* (e.g., deliberately allowing a disease to progress in a patient in order to test a treatment). Thus while both the DDA and the DDE tend to discriminate against actions in the first category and in favor of actions in the third category, the DDE may favor some actions in the second category that are not favored by the DDA, and the DDE may disfavor some actions in the fourth category that are not disfavored by the DDA.

⁷ The example that follows is based on Bennett's critical discussion of the DDE, though it is important to note that he is not endorsing the plausibility of the bomber's claims here: he agrees that they would be absurd. The problem he is raising is that if we cannot find a principled way to rule them out then the double effect distinction is too vague and arbitrary to do the moral work required of it by the DDE.

⁸ Masek does not claim that the live organ harvesting is ultimately *permissible*, but only that its impermissibility does not lie in the agent's intending the subject's death (since the death is unintended on his view), but instead lies in the fact that the organs are being stolen, for example. Thus, while his view doesn't entail an implausible moral verdict in such cases, provided other moral principles can explain their impermissibility, it does mean that in such cases the double effect distinction *applies yet fails to do the sort of justificatory work highlighted by the DDE*. As noted earlier, this needn't itself be a problem: the distinction needn't always do justificatory work where it applies. But multiplying such cases in this way only exacerbates the problem of explaining why it sometimes makes a moral difference and sometimes does not, putting most of the real explanatory burden on that account. Moreover, it has the implausible implication that the moral explanation of what is wrong with the vivisection-for-organs case is importantly different from what is wrong with the similar case where the subject is killed first as a means to get the organs (call it 'killing-for-organs'): the DDE's prohibition against intending the death of innocents is (for defenders of the DDE) a crucial part of the explanation of what's wrong with killing-for-organs, yet on Masek's view it is entirely absent from the account of what's wrong with vivisection-for-organs. Most defenders of the DDE, however, would see the two cases as being wrong for the same fundamental reasons, both involving intending the death of the subject as a means (among other things). If the central concern of the DDE with intending the death of innocents as a means or end were as easily deflected as it would be on Masek's view, it is unclear how it can at the same time be thought so critically important in the cases where it does do crucial moral work.

⁹ Thus Cavanaugh (2006, 66–72), for example, criticizes Finnis, Grisez, and Boyle (2001) for taking a line similar to Masek's above, applying the double effect distinction even to the craniotomy case. They claim that a doctor may legitimately claim not to intend the fetus's *death* (which is instead a foreseen but unintended side-effect) even as he acknowledges intentionally tearing apart the fetus's skull and its contents in the course of the procedure, as a means to saving the woman. Anscombe (1981, 2001) and Foot (1978), like Cavanaugh, reject applications of the intend/foresee distinction that allow such claims, seeing them as sophistical.

¹⁰ I here set aside the possibility of ignorance of the constitutive relation, which won't arise in cases like this. But in principle, if someone were truly (and bizarrely) ignorant of the fact that someone's being blown to bits is constitutive of his being killed, then she could intend the first without intending the second.

¹¹ The Trolley Case comes originally from Foot (1978), and variations on it will come up again later.

¹² At one point, Thomas Nagel (1979) discarded the DDE in favor of another deontological principle precisely because of worries about unclarity and arbitrariness in the application of the double effect distinction, particularly in connection with real and important circumstances in war (such as area bombing). He later came back to the DDE as a central part of his non-consequentialist view (Nagel 1986).

¹³ Quinn allows, of course, for exceptions, as in cases where one has a right to involve the person in this way even without his consent.

¹⁴ In the present context, the Trolley Case in question is not quite Foot's original, where you are the *driver* of the trolley, but a slight variant introduced by Thomson, where you are a *bystander* near a switch. The relevant comparison is now between the bystander version of Trolley and the Loop variant described in what follows.

¹⁵ Masek (2010) also defends the moral significance of the triple effect distinction. Note that Kamm herself, however, does not ultimately endorse the DTE as a fully adequate principle. Her own favored principle is different.

¹⁶ The next two paragraphs draw closely from FitzPatrick (forthcoming).

¹⁷ At least assuming, as we are for this case, that this pilot's bad intentions will not change his actions in any way: he will fly exactly the same mission as any other pilot, but will simply have different intentions and feelings in doing so. If his bad intentions were such as to undermine his reliability in carrying out the tactical raid according to his orders (e.g., making him likely to change plans mid-flight and target a city instead), then Scanlon would allow that it may be impermissible for him to fly the mission and for his commander to send him on it. But here the role of intention in affecting permissibility is *derivative*: it comes in merely in making the pilot unreliable and dangerous, which is what directly makes it impermissible for him to fly the mission (Scanlon 2008, 31–2). This is the focus of Alec Walen's (2006) "doctrine of illicit intentions," which, following Scanlon, rejects any direct or non-derivative role for intentions in affecting permissibility, as asserted by the DDE, but allows for this sort of indirect or derivative role.

¹⁸ Scanlon puts this point in terms of distinguishing *critical* and *deliberative* uses of moral principles: only the critical use of principles, as when criticizing an agent for having acted badly by *failing to act on the right considerations*, is essentially sensitive to the agent's state of mind; by contrast, the deliberative use of principles, as in judging permissibility, is essentially concerned simply with *what counts for or against potential acts*: for it is simply such considerations that *make* an act permissible or not (2008, 23).

¹⁹ Compare Thomson (1999, 515–6): "According to [the DDE, if combined with the assumption that it is always wrong for a doctor to intend the death of a patient], the question whether it is morally permissible for the doctor to inject a lethal drug turns on whether the doctor would be doing so intending death or only intending relief from

pain. That is just as absurd an idea. If the only available doctor would inject to cause the patient's death, or is incapable of becoming clear enough about her own intentions to conclude that what she intends is *only* to relieve the patient's pain, then—according to PDE—the doctor may not proceed, and the patient must therefore continue to suffer. That cannot be right.”

²⁰ Matthew Hanser (2005) has drawn a helpful distinction between *adjectival* and *adverbial* permissibility questions and judgments, i.e., between the question whether an act (however described) is *permissible* and the question whether an agent has *acted permissibly* in doing what she has done. The former concerns an act type while the latter is about the agent's concrete action. While the DDE is primarily concerned with the former notion, if one instead focuses on the latter and understands the notion of acting permissibly to be closely related to that of acting well, such that one acts impermissibly when one acts on bad or at least insufficient reasons, then this would provide another way of linking actual intentions to permissibility (since the bad-intentioned tactical bomber, though he does what is *permissible*, would be *acting impermissibly* due to his acting on the wrong reasons).

²¹ It is consistent with this approach to grant part of the earlier claim about thick act-types, allowing a very limited and qualified role for actual intentions in affecting permissibility. Since it's true that instantiating the thick act-type 'bombing the munitions plant with the intention of killing civilians' would constitute acting badly, and there is pro tanto reason to avoid acting badly, there is pro tanto reason for the bad-intentioned pilot to avoid flying the mission, instead finding a substitute if one is available; if there is a substitute available, then it might therefore be impermissible for this pilot to go ahead and fly it himself, thus needlessly acting badly. But if no other pilot is available and the mission is vitally important, then on the current account it may be not only morally *permissible* for him to fly the mission, but even morally *required*, despite the fact that he will be acting badly in doing so given his bad intentions and motives. That is, he will do something *permissible* even in performing the problematic act-type, though he will act badly in doing so (just as anyone does who performs a permissible action with bad motives, for example). This account thus has somewhat different implications from the previous account that appeals to thin vs. thick act-types as the primary answer to the objection from Rachels, Thomson and Scanlon (FitzPatrick forthcoming).

Works Cited

- Anscombe, G. E. M. *Ethics, Religion and Politics*. Minneapolis: University of Minnesota Press, 1981.
- . 'Medalist's Address: Action, Intention, and Double Effect.' *The Doctrine of Double Effect*. Ed. P. A. Woodward. Notre Dame: University of Notre Dame Press, 2001. 50–66.
- Bennett, J. *Morality and Consequences*, The Tanner Lectures on Human Values II, 45–116. Salt Lake City: University of Utah Press, 1981.
- . *The Act Itself*. Oxford: Oxford University Press, 1995.
- Cavanaugh, T. A. *Double-Effect Reasoning: Doing Good and Avoiding Evil*. Oxford: Oxford University Press, 2006.
- Finnis, J., G. Grisez and J. Boyle. "Direct" and "Indirect": A Reply to Critics of Our Action Theory.' *The Thomist* 65 (2001): 1–44.
- FitzPatrick, W. J. 'Acts, Intentions and Moral Permissibility: In Defense of the Doctrine of Double Effect.' *Analysis* 63.4 (2003): 317–21.
- . 'The Intend Foresee Distinction and the Problem of 'Closeness.' *Philosophical Studies* 128 (2006): 585–617.
- . 'Intention, Permissibility and Double Effect.' *Oxford Studies in Normative Ethics* 2 (forthcoming).
- Foot, P. 'The Problem of Abortion and the Doctrine of the Double Effect.' reprinted in her *Virtues and Vices*. Berkeley: University of California Press, 1978: 19–32.
- . 'Killing and Letting Die.' *Killing and Letting Die*. Eds. B. Steinbock, A. Norcross, 2nd ed. New York: Fordham University Press, 1994. 280–9.
- Hanser, M. 'Permissibility and Practical Inference.' *Ethics* 115.3 (2005): 443–70.
- Hills, A. 'Intentions, Foreseen Consequences and the Doctrine of Double Effect.' *Philosophical Studies* 133 (2007): 257–83.
- Kamm, F. M. *Morality, Mortality*, vol. II. Oxford: Oxford University Press, 1996.
- . *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. Oxford: Oxford University Press, 2007.
- Masek, L. 'Intentions, Motives and the Doctrine of Double Effect.' *The Philosophical Quarterly* 60.240 (2010): 567–85.
- McMahan, J. 'Intention, Permissibility, Terrorism, and War.' *Philosophical Perspectives* 23 (2009): 345–72.
- Nagel, T. 'War and Massacre.' in his *Mortal Questions*. Cambridge: Cambridge University Press, 1979. 53–74.
- . *The View From Nowhere*. Oxford: Oxford University Press, 1986.
- Otsuka, M. 'Double Effect, Triple Effect and the Trolley Problem: Squaring the Circle in Looping Cases.' *Utilitas* 20.1 (2008). 92–110.
- Quinn, W. 'Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing.' reprinted in his *Morality and Action*. Cambridge: Cambridge University Press, 1993a. 149–74.
- Quinn, W. 'Actions, Intentions, and Consequences: The Doctrine of Double Effect.' reprinted in his *Morality and Action*. Cambridge: Cambridge University Press, 1993b. 175–93.
- Rachels, J. 'Active and Passive Euthanasia.' *New England Journal of Medicine* 292 (1975): 78–80.

- . ‘More Impertinent Distinctions and a Defense of Active Euthanasia.’ *Killing and Letting Die*, 2nd ed. Eds. B. Steinbock, A. Norcross. New York: Fordham University Press, 1994. 139–54.
- Scanlon, T. M. *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge: Belknap, Harvard University Press. 2008.
- Thomson, J. J. ‘The Trolley Problem.’ *The Yale Law Journal* 94 (1985): 1395–415.
- . ‘Physician-Assisted Suicide: Two Moral Arguments.’ *Ethics* 109 (1999): 497–518.
- Walen, A. ‘The Doctrine of Illicit Intentions.’ *Philosophy and Public Affairs* 34.1 (2006): 39–67.
- Wedgwood, R. ‘Scanlon on Double Effect.’ *Philosophy and Phenomenological Research* 83.2 (2011): 464–72.