

The Doctrine of Double Effect as an Objective Principle

Howard Nye
University of Michigan

Abstract: The Doctrine of Double Effect states roughly that there are stronger moral reasons against inflicting harm as a means to an end than there are against causing harm as a byproduct. On one reading this is a claim about reasons not to act with certain intentions, while on another it speaks against performing acts with a certain objective causal structure. While there are good reasons to prefer the objective reading, some have dismissed it as absurd. In this paper I defend the objective reading against this dismissal, which seems largely to depend upon the conflation of criteria of objective wrongness with other criteria. I present an objective formulation of the DDE according to which a harm's bringing about a particular good weakens the extent to which that good counts in favor of bringing about the harm. This solves several problems and captures an intuitive injunction against benefitting at others' expense.

The Doctrine of Double Effect as an Objective Principle

Howard Nye
University of Michigan

1. Two Readings of Double Effect

The Doctrine of Double Effect [DDE] states roughly that:

DDE₁: There are stronger moral reasons against bringing about harm as a means to an end than against bringing about harm as a byproduct

The DDE is invoked to explain such non-consequentialist¹ intuitions as why:

(1) it seems permissible to divert a trolley from hitting five people to hitting one, but wrong to push someone into the trolley's path to stop it from hitting five others,

(2) it seems permissible to bomb enemy military targets and kill civilians as collateral damage, but wrong to kill civilians as a means of demoralizing the enemy into surrender, and

(3) it seems permissible to withhold medicine from one to give it to five others, but wrong to withhold medicine from one to make his organs available for harvest.

There are at least two ways of reading the DDE. The first interprets it as a claim about reasons not to act with certain mental states:

DDE_M: There are stronger moral reasons against intending to bring about harm than against doing things one merely foresees will result in harm.

¹ Or more accurately intuitions that go against *direct* and *universalistic* consequentialism, of which Direct Utilitarianism is perhaps the most common example.

The second interprets the DDE as a claim about reasons not to perform acts with a certain causal structure:

DDE_C: There are stronger moral reasons against bringing about good as a result of bringing about harm than against bringing about harm as a byproduct of bringing about good.

Against DDE_M, many have argued powerfully that an act's moral permissibility cannot depend upon the intentions with which it is performed.² At the same time, many have argued that anything like DDE_C is absurd.³

In this paper, I argue against the second group of critics that the DDE is best interpreted as a principle about objective causal structure. Some of the case against this reading may trade on implausible interpretations of the DDE which are orthogonal to the issue at hand. For this reason I begin with some ground clearing to show how a version of the DDE could be plausible. I then present the theoretical case against DDE_M, and respond to dismissals of DDE_C that result from conflating criteria of objective wrongness with other criteria. I proceed to identify genuine problems with the DDE_C formulation, but present a better interpretation of the DDE that solves these problems and brings out the DDE's plausibility.

2. Ground Clearing

While our examples (1), (2), and (3) all involve lethal harms, it seems clear that the DDE retains its force against non-lethal harms – e.g. torturing one person to obtain information necessary to prevent the torture of three others. The DDE's general concern with harm rather than death *per*

² See e.g. Gibbard (1990), Bennett (1995), Thomson (1999), Scanlon (2000), and Kamm (2008).

³ See e.g. Frey (1975), Bennett (1995), and Norcross (1999, 2008).

se entails that absolutist formulations of the doctrine are completely implausible. While heightened reasons against harming as a means might make it wrong to step on one foot to prevent two foot steppings,⁴ these heightened reasons would clearly be outweighed if one had to step on a foot to save a life.

Moreover, the DDE should be understood as claiming that there are heightened reasons against harming some beings to help *other* beings. If we must amputate Jones's foot to make a serum that will save his life, then so long as he does not object (say he is unconscious with no known preferences), there seems nothing wrong with doing so.⁵ But amputating the unconscious Jones's foot to obtain serum to save Smith seems to be exactly the kind of conduct against which there are heightened reasons.

Since on a proper understanding the DDE attributes heightened reasons against inflicting harm, it will not speak against killing someone as a means when death is good for him.⁶ Quinn (1989) and McMahan (2002) are thus clearly correct that principles like the DDE are misapplied in attempts to discriminate between using morphine to control pain with death as a side effect and using it to end suffering by ending life.

Finally, the DDE should be understood to speak against harm as a means only when all else is held equal – plausible defeaters include consent to harm, obligation to incur harm, and the culpability of those harmed.

⁴ This kind of example derives from Darwall (2006).

⁵ This case is an adaptation of versions of Unger's (1996).

⁶ Moreover, since we saw the DDE does not speak against harming a being for his own greater good, this will be so even if death is harmful in a respect but not on the whole.

3. Motivating Double Effect

I think these observations about plausible interpretations of the DDE help reveal problems with standard attempts to motivate the Doctrine in terms of platitudes against “using,” “exploiting,” or “treating agents as means.”⁷ These platitudes have little to do with harming a being on the whole. “He was just using me” seems most apt as a complaint about being benefitted as a means to a manipulator’s ends. Similarly, complaints about exploitation are usually raised about making people better off than they would have been, but in ways that take advantage of their inferior power.⁸

Moreover, as Bennett (1995) has argued, platitudes about using someone as a means cannot explain the DDE’s application to cases of harms that eliminate obstacles.⁹ The DDE appears to be exactly what explains why it seems wrong to do things like fatally blast a fat man (who would otherwise be fine) out of a cave mouth to save five others from drowning inside.¹⁰ But there is no clear sense in which this treats him as a useful instrument – oh, that he hadn’t been around at all!

While I think platitudes about use, exploitation, and means fail to motivate the DDE, I believe that we can locate its intuitive appeal elsewhere, with platitudes about our reasons against benefitting at the expense of others. There is something disturbingly predatory or ghoulish about accruing advantages from others’ misery, which comes out in how we think about sadistic behavior, thrill-killings, rape, taunting the defeated, aggressive cannibalism, and monstrous fictional agents (e.g. vampires). Problematic forms of benefit at others’ expense are

⁷ For motivations of this kind see e.g. Quinn (1989), Kamm (2007), and Richardson (2008).

⁸ This seems true of such complaints about the treatment of impoverished workers, the patronization of prostitutes and other sex-workers, and perhaps most vividly objections to patronizing a desperate “whipping boy for hire” who offers to allow you to beat him in exchange for his subsistence.

⁹ Indeed Quinn (1989) - the arch-user of such platitudes - himself came close to admitting this.

¹⁰ Discussion of such a case traces at least to Foot (1967).

often by means of elimination, including killing relatives for inheritance money and killing or framing people so one or one's friends can take their jobs or partners.

4. Moral Reasons and Intentions

The intuitions we have seen in favor of the DDE concern moral reasons for and against performing acts and how these reasons stack up to make acts morally wrong or permissible.

There is such a thing as doing what one should do or what one is morally permitted to do for bad reasons – e.g. if one refrains from cheating a child merely to avoid getting caught, or diverts a trolley from five to one out of spite for the one.

Questions about what to do and what is morally permissible are forward-looking and action guiding. When we come up with answers to these questions, we generally act on our views, intending to do what we think we should and intending not to do what we think we shouldn't. When we believe that M is a causally necessary means of doing what we should, we will generally intend to bring about M as a means. When we merely believe that B will arise as a side-effect of doing as we should, we will foresee that B will occur without intending to bring it about.¹¹ Thus, if we believe that DDE_C correctly describes our reasons, we will tend to act on it, be more reluctant to bring about good by bringing about harm, and thus be more reluctant to intend harm as a means in our pursuit of good.

While questions about our intentions in doing A do not bear on questions about our moral reasons to do A,¹² they are highly relevant to questions about which normative views are guiding

¹¹ For instance, as Frey (2003) points out, when we think we should wear our clothes to keep warm or drive our cars to get to work, we can foresee that our activities will wear out our clothes and cars without our intending to wear them out. This is because we do not take the wearing out of our clothes and cars to play any causal role in our accomplishing our ends of keeping warm and getting to work.

¹² As Bennett (1995) and Scanlon (2000) point out, we could have views about our reasons that make a kind of reference to our intentions. Scanlon notes that we might think we should not continue a conversation with someone about her running for an office we ourselves intend to run for, and Bennett notes that we might take the fact that we

our conduct. Unlike assessments of moral wrongness and permissibility, assessments of moral blameworthiness and estimability look back at the reasons on which an agent acted and, among other things, hold failures to act on genuine moral reasons against her.¹³ Thus, the plausibility of something like DDE_C as a criterion of right would explain that of DDE_M as a principle for assessing blame and disestimability. DDE_M's popularity as a criterion of right seems to result from confusing permissibility and wrongness with estimability and blameworthiness.

5. Norcross and the Pope's Reaction

The DDE is primarily supposed to be an action-guiding criterion of right that helps us determine what to do when faced with options to divert trolleys, push fat men, drop bombs, withhold medicines, etc. The foregoing thus supports DDE_C over DDE_M as a reading of the DDE. But Bennett (1995) simply dismisses DDE_C as "quite implausible," citing only a discussion of Frey (1975). Much of Frey's discussion turns on independently implausible features of historical versions of the DDE which I have dispensed with in my ground clearing. But I think Norcross's (1999, 2008) discussion of his Meteor case captures the essence of Frey and Bennett's idea without these distractions.

Norcross has us imagine a case in which Moe does what his evidence suggests will save five people from a meteor but will kill Homer as a side effect. Having the same evidence as Moe, the Pope awards Moe a Seal of Approval. But it turns out that Moe's evidence about a triggering device was misleading, and his act actually saved the five as a causal consequence of

could not help ourselves from doing something intentionally to be a reason against doing it at all. But these are not cases in which we actually think it morally impermissible to have certain intentions – in these cases we rather think morality requires us to do certain things if we have certain intentions or to bring it about that we will not be acting out of certain intentions.

¹³ Or do so when these failures are responsible and in a way that accounts for difficult or exculpatory conditions – see below.

killing Homer. When the Pope learns of this, she indignantly snatches the Seal of Approval from Moe. After telling this story Norcross concludes, “That the mechanics of a triggering device could make a moral difference, without any further difference in terms of who lives or dies, or in terms of [Moe’s] mental states, is too implausible to merit further discussion.”

The first thing to observe is that Norcross’s dismissal would be implausible in the absence of his story about Moe and the Pope. Most non-consequentialist theories countenance the relevance of factors other than who lives and who intends what. Many would find it highly plausible that the mechanics of a device could make a difference to one’s moral reasons if, for instance, they determined whether one’s pressing a button diverted a trolley from five to one or dropped a fat man into the trolley’s path.

But Norcross’s story is nothing but an attempt to get us to conflate criteria of objective wrongness with criteria of subjective wrongness, blameworthiness, and estimability. Throughout history, act-consequentialists have themselves been unfairly treated by similar confusions, and quite frankly they should know better. Consider the following counter-story:

Lisa does what her evidence tells her is saving many, and Mr. Burns does what his evidence tells him is harming many. Having the same evidence as Lisa and Mr. Burns, Norcross awards Lisa the Norcross Medal of Consequence. But it turns out that Lisa’s and Mr. Burns’s evidence is misleading – one of the poor people Lisa saved inadvertently tripped on a rock, which unleashed a disease from Springfield Gorge, wiping out most of Springfield. The only reason the disease didn’t wipe out neighboring Shelbyville too is that Mr. Burns shot the carriers of the disease for pure sport. When Gordon Gecko hears of this, he angrily strips Lisa of the Norcross Medal and gives it to Mr. Burns.

As has been pointed out repeatedly since (Mill 1861), the inappropriateness of Gecko's reaction constitutes absolutely no evidence against act consequentialism's identification of right acts with those that promote the best consequences. Consequentialism gives us a criterion of objective rightness, which we are to try to approximate by using our evidence to determine the likelihood that our acts will have the properties of which the theory speaks. As we lack omniscience, these subjective assessments of wrongness and reasons in light of our evidence are all we can be guided by in practice.

Whether an agent behaved rightly in the objective but not the subjective sense is completely irrelevant to the assessments of blame and esteem that are tied to rewards and punishments like medals and the snatchings of medals from necks. To be morally blameworthy, one must do what is wrong given one's evidence and face no further exculpatory circumstances like bad moral training, psychological disturbance, etc.¹⁴ Moral estimability is also a matter of doing what is supported by one's moral reasons in the subjective sense, in addition to doing it responsibly and for the right reasons.

DDE_C, like consequentialism formulated in terms of actual consequences, is an objective theory of moral reasons. It is understood that of course it will be implemented as a subjective theory through rational expectations of which courses of action will end up inflicting harm as a means. Since Moe applied his best evidence to determine the dictates of DDE_C in his circumstances, proponents of DDE_C will view him as reasoning exactly as he should, and doing right in the subjective sense. Moreover, as Moe was motivated by what he took to be his moral reasons in light of DDE_C and his evidence, he will be a good candidate for moral praise if DDE_C

¹⁴ See e.g. Gibbard (1990) for the general distinction and the example of emotional disturbances like a paroxysm of grief as an example of an exculpatory circumstance.

is true. The Pope's absurd failure to assign praise and blame in accordance with these standards is no more evidence against DDE_C than Gecko's is evidence against consequentialism.

6. Double Effect as a Weakening Principle

There remain, however, some serious problems with the DDE_C formulation of the DDE. DDE_C seems to claim that there are stronger reasons against bringing about a given harm if that harm will result in good. But surely this is absurd – why should the fact that good will come of a harm make it worse to bring it about?¹⁵

Because of this, DDE_C seems to have other unacceptable consequences. Consider a pair of cases developed by McMahan (1994).¹⁶ In the first, an accident victim will die if you don't help him, but there is a high probability of your contracting a deadly disease if you help. If the risk is significant enough, it will be supererogatory but not obligatory to help the victim. But now consider a version of the case that is identical except you also know that if you fail to help the victim, his organs will be used to save five others in hospital. DDE_C seems to say that there is actually stronger reason to help the victim in the second case than in the first – and that the mere addition of the five in hospital could make it obligatory to provide help that it would be supererogatory to provide otherwise!

Next, consider a case developed by Kamm (2008) following Norcross (2008). A trolley will run over five people on a main track unless you divert it to a side track, where it will run over one person. But if it runs over the one person, it will also pop off his head, which will roll in front of a tractor, and stop it from running over five completely different people. Intuitively it remains just as permissible to divert the trolley in this case as in (1) from Section 1. But DDE_C

¹⁵ See e.g. Norcross (2008) for related criticism.

¹⁶ The idea of something like these was suggested by Quinn (1989). In effect I present these cases in an order that is opposite McMahan's (though his versions involve differences in intention rather than objective causal structure).

seems to attribute the same heightened reasons against diverting in this case as against pushing the fat man in (1) – because the harm to the one will do the same amount of good!

To see what has gone wrong with DDE_C , we must attend to Dancy's (2004) distinction between considerations that oppose acts and considerations that weaken other reasons. For instance, while the fact that one has promised to do A is a reason to do A, the fact that the promise was given under duress is a consideration that weakens this reason without itself counting for or against doing A.

Intuitively, the fact that an accident victim's death will be a means to saving five others doesn't give us a new reason *to help* the accident victim. Rather, the fact that

S: Not helping him will save five others

is a powerful reason *against* helping. But what the idea behind the DDE - the injunction against benefitting some at the expense of others - seem to suggest, is that the fact that the good to the five will come through the death of the one *weakens* the extent to which S counts against saving the victim.¹⁷ This supports the following objective reading of the DDE:

DDE_F: All else held equal, the fact that an act will bring about harms that play a necessary causal role in bringing about a good effect, G, weakens the extent to which the fact that the act will bring about G counts in favor of performing the act.

¹⁷ Absolutist versions of the DDE would want to say that this completely disables S from counting against saving him. But as we have seen absolutist versions of the DDE are intolerable. We should admit that S really does retain some of its force against saving the victim, but not enough force to make it permissible to fail to save him in the absence of the fact that it is highly personally risky to do so. If the victim's death would instead result in the salvation of a bazillion people, it would surely be permissible not to save him absent the personal risk, and in fact obligatory not to save him either way.

DDE_F attributes no new reasons to help the accident victim when the five are in hospital, and so will not turn such supererogatory help into obligatory help.

The fact that diverting a trolley from five to one will leave the one dead is a strong reason against diverting, but too weak to outweigh our stronger reasons to save the five. This would also be true of pushing a fat man between a trolley and the five, except that by DDE_F the fact that their salvation would come by the fat man's death weakens the extent to which their salvation counts in favor of pushing him. In such a case, the fact that pushing will leave the fat man dead retains its full strength as a powerful reason against pushing, and is sufficient to decisively outweigh the now weakened reasons in favor of pushing.

In Kamm's case above, the fact that the death of the one plays a necessary role in saving the second five weakens the extent to which their salvation counts in favor of diverting. But in this case there is an entirely different reason to divert – namely that it will save the five on the main track. This reason retains its full strength and (just like above) is sufficient to outweigh one's reasons not to divert so as not to harm the one.

DDE_F delivers these results because it does not make the bizarre claim that it's worse to do harm if it does any good. It claims rather that a harm's bringing about a particular good weakens the extent to which that good counts in favor of bringing about the harm. This, I think, captures the intuitive injunction against benefitting some at the expense of others, which as far as I can see remains quite plausible.

REFERENCES

- Bennett, Jonathan. 1995. *The Act Itself*. Oxford: Clarendon Press.
- Dancy, Jonathan. 2004. *Ethics Without Principles*. Oxford: Clarendon Press.
- Darwall, Stephen L. 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge: Harvard University Press.
- Foot, Philippa. 1967. The Problem of Abortion and the Doctrine of Double Effect. *Oxford Review*, 5: 5-15.
- Frey, R.G. 1975. Some Aspects to the Doctrine of Double Effect. *Canadian Journal of Philosophy*, 5: 259-283.
- Frey, R.G. 2003. The Doctrine of Double Effect. In R. G. Frey, C. Wellman, *The Companion to Applied Ethics*. Oxford and Boston: Basil Blackwell, 464-474.
- Gibbard, Allan. 1990. *Wise Choices, Apt Feelings*. Cambridge, Massachusetts: Harvard University Press.
- Kamm, Francis. 2007. *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. New York: Oxford University Press.
- Kamm, Francis. 2008. Responses to Commentators on Intricate Ethics. *Utilitas*, 20: 111-142.
- McMahan, Jeff. 1994. Revising the Doctrine of Double Effect. *Journal of Applied Philosophy*, 11: 201-212.
- McMahan, Jeff. 2002. *The Ethics of Killing: Problems at the Margins of Life*. New York: Oxford University Press.
- Mill, John Stuart. 1863. *Utilitarianism*. London: Parker, Son, and Bourn.
- Norcross, Alastair A. 1999. Intending and Foreseeing Death: Potholes on the Road to Hell. *Southwest Philosophy Review*, 15: 115-123.
- Norcross, Alastair A. 2008. Off her Trolley? Francis Kamm and the Metaphysics of Morality. *Utilitas*, 20: 65-80.
- Quinn, Warren S. 1989. Actions, Intentions, and Consequences: The Doctrine of Double Effect. *Philosophy and Public Affairs*, 18: 334-351.
- Scanlon, T.M. 2000. Intention and Permissibility. *Proceedings of the Aristotelian Society*, 74 (Suppl): 301-317.

Thomson, Judith J. 1999. Physician-Assisted Suicide: Two Moral Arguments. *Ethics*, 109: 497-518.

Unger, Peter. 1996. *Living High and Letting Die: Our Illusion of Innocence*. New York: Oxford University Press.