Journal of Ethics and Social Philosophy
VOL. 12, NO. 3 · DECEMBER 2017
https://doi.org/10.26556/jesp.v12i3.267
© 2017 Author

IS AGENT-NEUTRAL DEONTOLOGY POSSIBLE?

Matthew Hammerton

It is commonly held that all deontological moral theories are agent-relative in the sense that they give each agent a special concern that *she* does not perform acts of a certain type rather than a general concern with the actions of all agents. Recently, Tom Dougherty has challenged this orthodoxy by arguing that agent-neutral deontology is possible.[1] His argument is simple: he posits a moral rule that he claims is both agent-neutral and deontological. In this article I show that the rule Dougherty posits cannot be both agent-neutral and deontological. The problem is that the rule has several possible interpretations and, although on some interpretations it is an agent-neutral rule, and on some interpretations it is a deontological rule, there are no interpretations in which it is both agent-neutral and deontological. I conclude by considering an alternative rule inspired by Dougherty's approach that looks like it might be both agent-neutral and deontological and showing that it too fails because it has core commitments that are incompatible with deontology.

1. DOUGHERTY'S ARGUMENT

Dougherty's argument starts from the assumption that deontological moral theories necessarily contain deontic constraints. Deontic constraints are moral rules that prohibit agents from performing acts of a certain type even if doing so is the only way to prevent more acts of that type from being performed by others. Thus, a deontic constraint on killing innocent people prohibits an agent from killing an innocent person even if doing so is the only way to prevent more killings of innocent people by others.

Deontic constraints have typically been thought to be agent-relative moral rules. For example, a deontic constraint on killing might be stated as follows:

KR: Each agent should ensure that she does not kill innocent people.

KR is an agent-relative rule because it gives each agent a special concern with acts

1 Dougherty, "Agent-Neutral Deontology."

of killing that she might perform, requiring her not to kill an innocent person even if doing so can prevent more killings by others. Following Parfit we can say that KR gives different aims to different agents.[2] For example, it gives me the aim that *I* do not kill and you the aim that *you* do not kill.

Dougherty suggests that although a deontic constraint on killing is normally formulated as an agent-relative rule like KR there is an agent-neutral alternative that could serve deontologists equally well:

> DR: Each agent should ensure that no one kills to prevent more killngs by others.[3]

Let us call the act of *killing to prevent more killing by others* "preventive killing." DR appears to be an agent-neutral rule because it gives all agents a general concern with preventive killings regardless of who they are performed by. Following Parfit we can say that it gives the same ultimate aim to all agents—the aim that there are no preventive killings.[4] DR also appears to produce a deontic constraint on killing because it prohibits each agent from killing an innocent person, even when doing so is the only way to prevent more killings by others. On these grounds Dougherty argues that a deontological theory could contain only agent-neutral rules like DR and thus that deontology can be agent-neutral.

Though this argument may appear plausible, I believe it is too fast. The problem is that DR can be given several different readings. To prove that agent-neutral deontology is possible there must be at least one reading of DR in which it is both agent-neutral and a deontic constraint. Let us consider each of the readings in turn.

First, consider a literal reading of DR. Read literally, DR requires each agent to make it the case that no one ever commits a preventive killing, and classes all agents as having violated DR and acted wrongly insofar as some agent, somewhere, kills preventively. To make this literal reading clearer, I will replace the word "ensure" with the word "guarantee":

> *Guarantee*: Each agent should guarantee that no one kills to prevent more killings by others.

Guarantee is an extremely demanding moral rule and, on these grounds alone, is implausible. Furthermore, *Guarantee* violates the *ought-implies-can* principle as it requires agents to ensure that there are no preventive killings even when they lack the ability and opportunity to do so. Many take the *ought-implies-can*

2 Parfit, *Reasons and Persons*, 27.

3 Dougherty, "Agent-Neutral Deontology," 531.

4 Parfit, *Reasons and Persons*, 27.

principle to be a conceptual truth, which would entail that *Guarantee* is not only implausible but also incoherent.[5]

These problems with *Guarantee* suggest that DR is better interpreted as containing a clause limiting it to cases in which an agent has the ability to satisfy it:

> *Ability*: Each agent, whenever she has the ability and the opportunity to do so, should ensure that no one kills to prevent more killings by others.[6]

Ability does not place unrealistic and excessive demands on agents and thus is much more plausible than *Guarantee*. However, *Ability* faces another problem. Consider the following case:

> *Prevention*: Tom and Mary are each about to commit a preventive killing. The only way for you to stop them is by committing one preventive killing yourself (perhaps the horror of witnessing a preventive killing will cause them to abandon their plans).

In *Prevention* you do not have the ability to ensure that no one commits a preventive killing. Thus, *Ability* does not class you as doing the morally wrong thing when you fail to stop all preventive killings. However, you do face a morally significant choice in *Prevention* between allowing two preventive killings to occur or committing one preventive killing yourself. *Ability* gives you no advice on what to do here. Yet a deontic constraint on killing must give you advice in this situation, requiring you not to perform a preventive killing even when doing so is the only way to prevent more preventive killings by others (as KR clearly does). So, although *Ability* appears to be agent-neutral, it is not a deontic constraint and thus fails to satisfy a necessary condition of deontology.

To address this problem we might further modify DR:

> *Oppose*: Each agent should, to the best of her ability, ensure that no one performs a preventive killing, making sure that whenever she cannot prevent all such killings, she at least does not perform any herself.

5 See Vranas, "I Ought, Therefore I Can," 167–69, for a defense of the "conceptual necessity" interpretation of the *ought-implies-can* principle (OIC). It should also be noted that those who reject OIC usually agree that the principle holds true for much of our moral theorizing but maintain that it fails in some special cases. However, *Guarantee* violates OIC in regular cases and not just in the special cases in which some argue that the principle fails.

6 Among the reasons why Dougherty might want to accept *Ability* is the fact that he borrows the term "ensure" from McNaughton and Rawling and apparently intends to use it in the same way that they use it. However, McNaughton and Rawling ("Value and Agent-Relative Reasons," 34) clarify that they interpret ensuring rules as containing a clause similar to that that appears in *Ability*.

Unlike *Ability*, *Oppose* does qualify as a deontic constraint on killing. For *Oppose* requires of each agent that she never kills, and requires this even if she is in a situation like *Prevention* in which, by killing, she can prevent more preventive killings overall. However, *Oppose* is not an agent-neutral rule. For the last clause in *Oppose* requires each agent to give special priority to her own killings, giving her the aim that *she* does not kill in cases like *Prevention*, over and above any general aim that no one kills. Thus, *Oppose* also fails as an example of agent-neutral deontology.

Finally, let us consider an alternative way of dealing with *Prevention* that preserves the agent-neutrality of DR:

> *Minimize*: Each agent should, to the best of her ability, ensure that no one performs a preventive killing, minimizing the total number of such killings whenever she cannot prevent them all.

Minimize appears to be agent-neutral as it gives all agents the same ultimate aim—the aim that there are as few preventive killings as possible. However, *Minimize* is not a deontic constraint as, in cases like *Prevention*, it does not, as a deontic constraint must, prohibit you from killing. Thus, *Minimize* also fails as an example of agent-neutral deontology.

I cannot see any other interpretations of DR that have the potential to give us a rule that is both a deontic constraint and agent-neutral. Thus, on the basis of the arguments above, I conclude that Dougherty's rule DR fails to demonstrate that deontology can be agent-neutral.

2. AN ALTERNATIVE RULE

The main problem with DR was that it cannot be an agent-neutral rule and yet also give a deontological verdict in cases like *Prevention*. Dougherty might try to fix this problem by postulating an additional agent-neutral rule that requires agents to ensure to the best of their ability that no one performs a second-order preventive killing (i.e., a preventive killing that itself prevents other preventive killings). However, this rule would be subject to a further counterexample, a case involving third-order preventive killings. In fact, for any additional rule concerning n-order preventive killing there is a *Prevention*-like counterexample involving $n + 1$-order preventive killings.

One way to escape this infinite sequence of counterexamples is to construct a rule requiring agents to prioritize stopping any higher-order killings over stopping any lower-order killings it prevents. To make this idea clearer let us say that when the absence of a higher-order killing is part of the causal explanation of why a lower-order killing occurs, then the higher-order killing is "causally up-

stream" of the lower-order killing. With this idea we can formulate the following rule:

> NR: Each agent should, to the best of her ability, ensure that outcome O1 occurs rather than outcome O2 whenever O2 contains a killing that is causally upstream of all the killings in O1.[7]

NR appears to be an agent-neutral rule as it gives all agents the same aim—that causally upstream killings are not performed. It also appears to give the same deontic verdicts as the agent-relative KR as an agent will always be prohibited from performing a preventive killing of any order because it will be causally upstream of any killings that will occur if she does not perform a preventive killing. Indeed, even killing to prevent several agents from violating NR is prohibited by NR because such a killing is causally upstream of the killings committed by those who are violating NR. Thus, NR appears to avoid the problems faced by DR.

Nonetheless, I think NR fails to capture an essential feature of deontology that agent-relative rules like KR are able to capture. To see this we need to consider the following case. Suppose that you and Ethel are both enemies of Joe and that each of you wants him dead. Furthermore, suppose you know that if Ethel has the opportunity to kill Joe she will and that, given your weak will, if you have the opportunity to kill Joe you will do so as well. Finally, suppose you know that either you or Ethel will eventually get an opportunity to kill Joe. Thus, an outcome in which neither of you kills him is not going to obtain. It follows that one of the following two outcomes will result:

> O1: You kill Joe.
> O2: Ethel kills Joe.

Importantly, neither of the killings in these two outcomes is causally upstream of the other. If O1 obtains, the correct causal explanation will appeal to the fact that you had the opportunity to kill Joe, the temptation to kill Joe, and a weak will unable to resist temptation. If O2 obtains the correct causal explanation will appeal to the fact that Ethel had the opportunity to kill Joe, the desire to kill Joe, and a morally corrupted will. In neither case will one of you kill Joe in order to prevent the other from killing him.

Now imagine in this scenario that you have the options of either ensuring that you do not kill Joe or ensuring that Ethel does not kill Joe but not the option of ensuring that neither of you kills Joe. For example, perhaps you and Ethel share the use of a car that tomorrow either of you might use to reach Joe and kill him. Tonight you must decide where to put the car keys. If you put them in one

7 This rule is adapted from a suggestion made by Tom Dougherty in private correspondence.

spot then Ethel will grab them first tomorrow and use the car to kill Joe. If you put them in the only other possible spot then you will have first access to them tomorrow and, given your weak will, you know that you will give in to temptation and take the car, using it to kill Joe. Thus, your choice of where to put the car keys is also a choice about whether to ensure that you do not kill Joe or ensure that Ethel does not kill Joe.

What should you do in such circumstances? I think it is clear what a deontological approach to killing must say here. It must say that in such circumstances your priority lies first and foremost with yourself. You must ensure that you do not kill innocent people. Maybe you are also required to ensure that others do not kill innocent people. However, if you have to choose between ensuring that you do not do this and ensuring that someone else does not do it you must (all else being equal) choose the former.

By requiring each agent to ensure that she does not kill, the agent-relative KR secures the correct deontological verdict in this case. However, NR is not able to secure this verdict. This is because neither O1 nor O2 is causally upstream of the other. Thus NR does not require the agent to ensure that O2 occurs rather than O1 (or vice versa). Instead, it gives the agent no guidance in this case. It appears then that agent-neutral rules like DR and NR are insufficient to produce moral theories that give the correct deontological verdicts across various cases involving deontic constraints. Only agent-relative rules like KR can produce those verdicts and thus the orthodox view that deontology is necessarily agent-relative is preserved.[8]

Singapore Management University
mhammerton@smu.edu.sg

REFERENCES

Dougherty, Tom. "Agent-Neutral Deontology." *Philosophical Studies* 163, no. 2 (March 2013): 527–37.

McNaughton, David, and Piers Rawling. "Value and Agent-Relative Reasons." *Utilitas* 7, no. 1 (May 1995): 31–47.

Parfit, Derek. *Reasons and Persons*. Oxford: Clarendon Press, 1984.

Vranas, Peter. "I Ought, Therefore I Can." *Philosophical Studies* 136, no. 2 (November 2007): 167–216.

8 I would like to thank Ryan Cox, Tom Dougherty, Daniel Nolan, and Nic Southwood for helpful comments on an earlier draft of this article.