CrossMark

Utilitas (2021), **33**, 474–479
doi:10.1017/S0953820821000121

CAMBRIDGE UNIVERSITY PRESS

REPLY

Setiya on Consequentialism and Constraints

Ryan Cox[1] and Matthew Hammerton[2]*

[1]School of Philosophical and Historical Inquiry, University of Sydney, Sydney, Australia and [2]School of Social Sciences, Singapore Management University, Singapore
*Corresponding author. Email: mhammerton@smu.edu.sg

Abstract
It is widely held that agent-neutral consequentialism is incompatible with deontic constraints. Recently, Kieran Setiya has challenged this orthodoxy by presenting a form of agent-neutral consequentialism that he claims can capture deontic constraints. In this reply, we argue against Setiya's proposal by pointing to features of deontic constraints that his account fails to capture.

1. Introduction

It is widely held that consequentialism, in its traditional agent-neutral form, is incompatible with deontic constraints and special duties.[1] Because deontic constraints and special duties are core features of common-sense morality, many take this incompatibility to be a serious problem for agent-neutral consequentialism. In 'Must Consequentialists Kill', Kieran Setiya challenges this widely held view.[2] He argues that there is a plausible version of agent-neutral consequentialism that can accommodate deontic constraints, but, at the same time, he concedes that special duties are, indeed, incompatible with agent-neutral consequentialism.

In this reply, we argue that Setiya's attempt to capture deontic constraints from within an agent-neutral framework fails. First, we argue that, even if his strategy works in the cases he discusses, there are other cases involving deontic constraints where it cannot work, and an agent-relative explanation of those constraints is unavoidable. Second, we show that the agent-relative explanation that is necessary in some cases of deontic constraints and special duties must be present in all cases. Thus, in any cases where agent-neutral reasons are adequate to explain a constraint, we must nonetheless accept that there are also agent-relative reasons present. A question then arises as to which reasons are doing the work in explaining the constraint.

[1]Deontic constraints are rules that prohibit performing an action even when doing so is the only way to prevent more instances of the action being performed by others. Special duties are rules requiring agents to favour those they have a special relationship with even when not favouring them would result in more special relationship favouring by others. Setiya calls the former 'general restrictions' and the latter 'special restrictions'. However, we prefer to use the more common terms 'deontic constraints' and 'special duties'.

[2]Kieran Setiya, Must Consequentialists Kill?, *The Journal of Philosophy*, 115 (2018), 92–105.

© The Author(s), 2021. Published by Cambridge University Press

https://doi.org/10.1017/S0953820821000121 Published online by Cambridge University Press

2. Consequentialising deontic constraints

Setiya begins his argument with a non-standard characterisation of agent-neutral consequentialism. To avoid difficulties arising from the notions of 'best' and 'good', as they occur in the statement of consequentialism, Setiya characterises consequentialism as the conjunction of the following two theses:

> ACTION-PREFERENCE NEXUS: Among the actions available to you, you should perform one of those whose consequences you should prefer to all the rest.
>
> AGENT-NEUTRALITY: Which consequences you should prefer is fixed by descriptions of consequences that make no indexical reference to you.

Setiya's characterisation is non-standard, because consequentialists often take notions like 'best' and 'good', as applied to consequences, to be utterly unproblematic. For the sake of argument, however, we will grant Setiya his non-standard characterisation of consequentialism (we will return to it below).

Setiya characterises the difference between preferences we ought to have which are agent-relative and those which are agent-neutral as follows: the preferences we ought to have are agent-relative just when the description of the consequences that the preference is about has an ineliminable indexical in it, and are agent-neutral otherwise. He later clarifies that when such an indexical is present, the preferences we ought to have are different for different agents, and, when it is absent, they are the same for all agents. This clarification coincides with Parfit's method for drawing the distinction.[3] To simplify things in what follows, we will characterise the distinction in this Parfitian way: the preferences we ought to have are agent-relative if they are different for different agents, and agent-neutral if they are the same for all agents.[4]

With this background in place, Setiya argues that deontic constraints *can* be accommodated by agent-neutral consequentialism, if only we are careful about how we apply the relevant tests for agent-neutrality and agent-relativity. Setiya asks us to consider a case which illustrates a familiar deontic constraint, namely, that constraint which prohibits killing an innocent person even if doing so is the only way to prevent more killings of innocent people by others:

> Murderous Footbridge: A villain has directed a trolley at five people in order to kill them and a bystander has the option of pushing an innocent person in front of the trolley, sacrificing this person to save the five.

The deontic constraint on killing prohibits sacrificing the innocent person in this case. It is commonly thought that only an agent-relative moral theory – a theory which gives different aims to different agents – can produce this verdict. However, Setiya rejects this

[3]Derek Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1984).

[4]We prefer to draw the distinction this way not only because it is easier to grasp than talk of ineliminable indexicals but also because, as McNaughton and Rawling have shown, the indexical approach is flawed (David McNaughton and Piers Rawling, Agent-Relativity and the Doing-Happening Distinction, *Philosophical Studies*, 63 (1991), 167–85). In certain cases it wrongly classifies agent-neutral rules as agent-relative (whereas Parfit's account gives the correct classification in these cases). Although Setiya's argument does not include any cases that the indexical account misclassifies, we nonetheless prefer to avoid using an account that has this well-known problem.

https://doi.org/10.1017/S0953820821000121 Published online by Cambridge University Press

common thought. He asks us to consider carefully what we ought to prefer to happen in Murderous Footbridge. Should we prefer that the one is sacrificed and the five are saved or that the one is spared, in which case the five do not escape murder? He suggests that, intuitively, everyone ought to prefer, in Murderous Footbridge, that the one is not sacrificed. The bystander who can make the sacrifice, a neutral third party observing the situation, the five victims who are going to be killed, and the villain who pushed the trolley, all ought to prefer, from a moral point of view, that the sacrifice is not made. If this is correct, then everyone ought to prefer the same thing in this case and thus the preference is agent-neutral. However, given ACTION-PREFERENCE NEXUS, this agent-neutral preference will result in a deontic constraint on killing. Thus, Setiya claims that the verdict of a deontic constraint can be produced by an agent-neutral consequentialist theory after all.

3. A problem with Setiya's account

So far so good. However, somewhat surprisingly, Setiya claims that his strategy for incorporating deontic constraints into agent-neutral consequentialism cannot be extended to special duties. This is surprising since, when it comes to agent-relativity, deontic constraints and special duties are commonly regarded as closely connected. It will be important for us to see why Setiya does not think that his strategy can be extended to the case of special duties for this will point us towards the difficulties for Setiya's account. To show that special duties are different, Setiya appeals to the following case:

> Future Neglect: Suppose that you and I both learn that in the future one of two things will happen. Either I will neglect my child, who will suffer terribly. Or, independently, you will neglect your child, who will suffer terribly.

Setiya suggests that it is intuitively obvious in this case that each of us ought to prefer that it is the other parent who is neglectful rather than ourselves. But if this is the case then we ought to prefer different things; I ought to prefer your neglect to mine and you ought to prefer my neglect to yours. Therefore, what we ought to prefer is agent-relative.

We agree with Setiya that this case demonstrates that special duties have an unavoidable agent-relative aspect that cannot be produced by agent-neutral consequentialism.[5] However, we think that Setiya is mistaken to think that this shows an important difference between special duties and deontic constraints.

The problem is that the lesson Setiya draws about special duties from Future Neglect can also be drawn about deontic constraints from other similar cases. For example, consider the following case:

> Indeterminate Trolley: John has sent a trolley hurtling towards his enemy in an unjust attempt to kill her. Mary has sent a trolley hurtling towards her enemy in an unjust attempt to kill him. If either trolley hits its targets it will succeed in killing them. An emergency system is operating that will detect any collision between a human and a trolley and then immediately shut down all remaining

[5]Actually, we find something puzzling in Future Neglect. If your neglect is independent of mine then shouldn't it also be possible for neither of us to neglect our child? Yet it is crucial to Setiya's example that this is impossible. Ultimately, we don't think this matters because there is a modified version of this case (a version that has the same structure as Indeterminate Trolley bellow) that isn't puzzling in this way and thus is sufficient to establish Setiya's point. On this basis we say we agree with Setiya here.

https://doi.org/10.1017/S0953820821000121 Published online by Cambridge University Press

> trolleys. Thus, if John's trolley reaches its target first, Mary's trolley will be disabled and her victim will be saved. Likewise, if Mary's trolley reaches its target first, John's trolley will be disabled and his victim will be saved. Now let's suppose that it is currently indeterminate which trolley will hit first.

Which outcome should John and Mary prefer in this case? Let's assume that each victim is equally deserving of being spared their fate. In that case, a neutral bystander ought to be indifferent. However, the same does not seem true of John and Mary. Even though the end result will be the same in terms of one person being unjustly killed, it does seem that John ought to prefer the outcome in which his attempted killing fails and Mary ought to prefer the outcome in which her attempted killing fails. On the face of it, this is an example of a deontic constraint which cannot be captured by Setiya's style of explanation.

In a footnote, Setiya briefly discusses a similar case in anticipation of this line of objection and suggests that the differences in preference can be accounted for on the basis of personal considerations (I would rather not be a murderer) rather than moral reasons related to a constraint on killing.[6] However, although there might be personal reasons for these preferences there also appear to be distinctly moral reasons. To see this, we can imagine that the conditions in Indeterminate Trolley are such that neither John nor Mary have any prudential reasons for preferring that the other's trolley hits first. For example, perhaps John knows that he will escape social censure and punishment if his murder attempt succeeds. And perhaps he knows that the guilt he will feel if his murder attempt succeeds is no greater than the guilt he would have felt for an attempted murder that failed only because Mary's murder attempt succeeded. Yet, even when we imagine that John lacks these kinds of prudential reasons, it still seems that he ought not be indifferent about which murder attempt succeeds. To be indifferent appears to miss something morally important about his specific connection to the killing that he may cause.

What could explain this difference in moral reasons? Two broad types of explanation for why agents ought to have a special moral concern with their own killings, as they seem to have in cases like Indeterminate Trolley, have been offered by moral philosophers.[7] An *agent-focused* explanation appeals to the agent's moral integrity. Separate from my prudential concerns, I have a moral concern with not being a person who does unjust things to others, and I have this concern even if keeping my hands clean will make no difference to whether the world that results is a better world overall. A *victim-focused* explanation appeals to a duty each agent owes to her potential victim not to harm her. My potential victim has a special claim on me that I do not harm her, and this claim outweighs the claim that someone else's potential victim has on me that her perpetrator not harm her. The result of this special claim is that, other things being equal, I ought to prefer that I do not harm my victim over someone else not harming their victim.

Whichever explanation is preferred, the upshot is that in Indeterminate Trolley different agents ought to prefer different things. Thus, there are cases involving deontic

[6]See Setiya, footnote 14.

[7]For influential accounts of these two kinds of explanations, see Thomas Nagel, *The View from Nowhere* (Oxford: Oxford University Press, 1986), pp. 175–85; F. M. Kamm, Harming Some to Save Others, *Philosophical Studies*, 57 (1989), 227–60; and F. M. Kamm, *Morality, Mortality: Volume II: Rights, Duties, and Status* (Oxford: Oxford University Press, 1996).

https://doi.org/10.1017/S0953820821000121 Published online by Cambridge University Press

constraints in which the preferences we ought to have are agent-relative. In this respect, deontic constraints are no different from special duties. Both have agent-relative aspects that cannot be produced by Setiya's version of agent-neutral consequentialism.

This argument shows that Setiya has not succeeded in giving an agent-neutral explanation of *all* deontic constraints. However, he could still claim that an agent-neutral explanation of deontic constraints is the correct explanation in at least some cases. And this would be enough to challenge the orthodoxy. For example, he might claim that even if an agent-neutral explanation of deontic constraints fails in Indeterminate Trolley it nonetheless provides the best explanation of Murderous Footbridge.

We find such a position untenable. The problem is that to pursue such an explanation, one would have to deny the presence of agent-relative reasons, or at least their explanatory relevance, in cases like Murderous Footbridge, while allowing for their explanatory relevance in other cases.

For example, consider Indeterminate Trolley again. We saw in this case that John has an agent-relative reason for preferring that he does not kill his potential victim. Now suppose that John finds himself in the position of the bystander in Murderous Footbridge. He can kill an innocent person to save the five that the villain has attempted to kill. Setiya shows us that there is an agent-neutral reason for John to prefer that he does not kill the one. However, in addition to this agent-neutral reason, doesn't John also have an agent-relative reason? Doesn't the agent-relative reason he had in Indeterminate Trolley also tell him to prefer that he doesn't kill in this case? If Setiya wants to claim that Murderous Footbridge has a purely agent-neutral explanation, then he must deny that the agent-relative reason carries over to this case. But this is implausible. There are no principled grounds for limiting agent-relative reasons in this way. Furthermore, the two main explanations given for agent-relative reasons – agent integrity and victim inviolability – are just as much applicable to Murderous Footbridge as they are to Indeterminate Trolley. Thus, at best, Murderous Footbridge can be understood as a case where an agent's reason for preferring something is overdetermined. On this understanding, the bystander has both agent-relative and agent-neutral reasons and each alone is sufficient to determine what he ought to prefer.[8]

4. The order of explanation

We will end our discussion by raising an issue about the order of explanation in the reasons we have been discussing. Setiya shows that there are agent-neutral reasons for preferring that agents abide by deontic constraints. We have shown that they must sit alongside agent-relative reasons. Arguably, the best explanation of the agent-neutral reasons Setiya appeals to is one that derives them from the agent-relative reasons. For example, consider Murderous Footbridge. An agent-relative reason explains why the bystander ought to prefer that he does not sacrifice the one and why it would be wrong for him to do this. Insofar as there is also an agent-neutral reason for preferring that he doesn't sacrifice the one, the following explanation looks attractive: everyone ought to prefer that the bystander does not sacrifice the one in this case because sacrificing the one is morally wrong. And sacrificing the one is morally wrong

[8]As we are about to argue, the agent-neutral reason the agent has in this case derives from the agent-relative reason they have. If so, the case should not be understood as one where an agent's reason for preferring something is overdetermined.

https://doi.org/10.1017/S0953820821000121 Published online by Cambridge University Press

because of an agent-relative reason that counts against it (e.g. the inviolability of the victim, or the moral integrity of the actor).

In general, we ought to prefer that people act as they are morally required to.[9] That a certain act is the right act for an agent to perform explains why others ought to prefer that he act that way. We suspect that this is one of the reasons why Setiya went for a non-standard formulation of consequentialism that uses preferability in place of betterness. Because it is a truism that everyone ought to prefer (from a moral perspective) that others act in morally appropriate ways, our intuitive judgements about what everyone ought to prefer in cases of sacrificing some to save others will match up with our intuitive judgements about whether it is permissible to sacrifice some to save others. By contrast, because our judgements about which states of affairs are best are (at least to some degree) independent of our intuitive judgements about which acts are right, these judgements don't always match up. Thus, although common-sense morality tells us that it is wrong to sacrifice the one to save the five in Murderous Footbridge, our intuitive judgements do not tell us that the outcome of not sacrificing the one is morally better than the outcome of saving the five (in fact, they seem to tell us the opposite). The plausibility of Setiya's analysis of Murderous Footbridge crucially depends on replacing judgements of what is better and worse with judgements of what we ought to prefer because the latter, but not the former, are derived from the deontic verdicts we already accept.

This observation casts Setiya's project in a new light. Consequentialism has traditionally been understood as the view that what we ought to do is determined and explained by the value of the consequences we can bring about. By replacing value with preferability, and reversing the order of explanation such that what each agent ought to do explains the evaluative ranking of states of affairs, Setiya is making a significant departure from traditional consequentialism. For some, this will raise questions about whether Setiya's proposal can plausibly be construed as 'consequentialism'. However, we do not want to enter into this murky linguistic debate. Our point is more straightforward. Deontic constraints require agent-relative reasons, and the agent-neutral reasons that Setiya points to, if they exist, are best explained as being derived from more basic agent-relative ones.[10]

Competing interests

The author(s) declare none

[9]A possible exception to this rule is the supererogation cases discussed in Theron Pummer, Whether and Where to Give, *Philosophy and Public Affairs*, 44 (2016), 77–95, and Joe Horton, The All or Nothing Problem, *Journal of Philosophy*, 114 (2017), 94–107. However, the cases that we discuss here (and those discussed by Setiya) are not cases of supererogation with the relevant structure. Indeed, they are all cases where the link between what an agent is required to do and what others should morally prefer that she does is highly plausible.

[10]We would like to thank Andrew Forcehimes and an anonymous referee from this journal for helpful comments on an earlier version of this article.

Cite this article: Cox R, Hammerton M (2021). Setiya on Consequentialism and Constraints. *Utilitas* **33**, 474–479. https://doi.org/10.1017/S0953820821000121

https://doi.org/10.1017/S0953820821000121 Published online by Cambridge University Press