Against Animal Replaceability: A Restriction on Consequences

Ricardo Miguel, University of Lisbon, Portugal

Abstract

Animal replaceability is supposed to be a feature of some consequentialist theories, like Utilitarianism. Roughly, an animal is replaceable if it is permissible to kill it because the disvalue thereby caused will be compensated by the value of a new animal's life. This is specially troubling since the conditions for such compensation seem easily attainable by improved forms of raising and killing animals. Thus, grounding a strong moral status of animals in such theories is somewhat compromised. As is, consequently, their position as an alternative to rights-based theories in animal ethics. Recognising this, some utilitarians tried to disassociate utilitarianism and replaceability. I will here add my voice to this project. However, instead of seeing the culprit in the usual suspects (hedonism, maximisation or the total view), I advance a new proposal. After identifying that the compensating value for a disvaluable action has to be its consequence, I present a restriction on consequences: consequences of sequences of actions cannot be consequences of the isolated actions in the sequences. Given this, the main argument is simple: killing an animal is permissible only if the the value of the new animal's life is a consequence of the killing; but this value is a consequence of a sequence of actions which involves the killing plus some additional actions; therefore, since, via the restriction, such value is not a consequence of the killing, it is irrelevant to its normative status. I then present two further motivations for the restriction: firstly, it prevents the value of conditional actions from trivially influencing the value of the actions on which they are conditional; secondly, it is useful – even if not a complete solution – to reply other objections to consequentialism: the accordion effect of action and the cluelessness problem. I finally consider a couple of objections.

I   The Replaceability Argument

This is what I take to be the best available version of the replaceability argument (RA):[1]

(1) If killing animals[2] whose future life would have a positive value will lead to the creation of other animals which would not exist otherwise and whose lives will have at least the same value as the one lost with the killed animals, then such killing is permissible.

(2) Killing *this* animal exemplifies the antecedent of (1).

---

[1] See Miguel 2016 for a contrast between this and two other versions of the argument.

[2] I use 'animals' to abbreviate 'non-human animals'. Although the RA may apply to humans too, I choose to focus on animal replaceability mainly because of its greater practical importance.

(3)   It is permissible to kill *this* animal.

A theory that implies this argument is not just incapable of a strong moral protection of animals' lives – like a right to life – but, in addition, allows killing them given conditions which seem rather easy to satisfy. Even if most animals presently raised for some purpose that implies killing them do not enjoy good lives and, as such, are not individuals which make (2) true, some present or improved forms of raising and killing animals may find support in the RA. For example, according to the RA, the conscientious small farmer who raises animals for food is killing them permissibly.

In general, moral theories that require promoting overall value above individual harms and benefits seem to imply the RA. As a standard example of such theories, Utilitarianism has been criticised for recognising an inadequate moral status of animals and, relatedly, for not grounding ethical vegetarianism. If this is sound, Utilitarianism falls behind competing views in animal ethics that tick those marks, like rights-based theories.

Nevertheless, I must note that some authors have defended non-standard utilitarian views that do not imply the RA, but I cannot assess their merits here. Instead, I will propose a novel way to cut the link between Utilitarianism and replaceability – one that restricts the notion of consequence and maintains the core properties of the standard theory.

II   Necessary Conditions for the Compensating Value

Suppose that some value, *v*, is not a consequence of some action, ɸ. Then *v* does not determine, or contributes to determine, the normative status of ɸ. Therefore, when a value compensates a disvalue which is a consequence of ɸ, the former must be a consequence of ɸ too. This means that the success of the RA implies that the value of the new animal's life – the candidate to compensating value – must be a consequence of killing another animal – the action with a disvaluable consequence.[3] However, as I will argue, there are good reasons not to regard such value as a consequence of the killing. But before moving on to this, let me illustrate why the consequentialist is committed with this tight relation between the compensating value and the value it is to compensate.

---

[3] This action may have, and normally has, valuable consequences too. Throughout I use 'disvaluable action' just to mean the action with the relevant disvaluable consequence, which is compatible with it being, sometimes, permissible.

In November 2017 a lynx that had escaped from an animal park in Wales was shot dead. Imagine that right after this a new lynx was born at the park and that his life was at least as valuable as the killed lynx's future life would be if he had not been killed. Thus, the balance between the value of the new lynx's life and that lost with the killing would not be negative. Yet, the value of the new lynx's life does not compensate the killing in the required sense aimed at by the RA – it does not make it permissible. Why? Since the new lynx's existence is independent of the other lynx's death, whatever value his life has, it is not a consequence of the killing. Therefore, such value cannot determine the normative status of the killing. In addition, this example brings to light that, when the disvaluable action is independent of the alleged compensating value, its omission would have made things better.

Thus, besides the requisite of non-negative net value, we have two other necessary conditions for the compensating value: on the one hand, it has to be a consequence of the disvaluable action (consequentialism); on the other hand, the omission of the disvaluable action and the performance of that which leads to the compensating value has to be inaccessible to the agent (maximisation). In sum, utilitarian value compensation requires: (i) that an action ψ brings about a value at least as good as the one lost with a disvaluable action ϕ; (ii) that the value of ψ is a consequence of ϕ; and (iii) that performing ϕ and ψ maximises the good.

To my knowledge, everyone discussing this matter has been accepting that the RA satisfies (ii).[4] I think that this is wrong and will argue for a restriction according to which (ii) fails.

III   A Restriction on Consequences

III.a   Blocking the Replaceability Argument

Consider the following restriction on consequences:

(R) Consequences of sequences of actions cannot be consequences of the isolated actions in the sequences.[5]

---

[4] To name a few, see Singer (2011), Regan (2004), Višak (2013, 2016), Chappell (2015) and Delon (2016).

[5] I am shamelessly applying to my needs Diogo Santos' "Non-disaggregation Principle" (ms.), which he uses to deal with the cluelessness problem (see the end of III.b). After reading Bratman (2006) on the connections between the accordion effect and Hart and Honoré's (1959) Voluntary Intervention Principle, I realised that (R) also has some connections with that principle, but I cannot explore them here.

I will now show that (R) blocks the RA. Recall premiss (1): If killing animals whose future life would have a positive value will *lead to* the creation of other animals which would not exist otherwise and whose lives will have at least the same value as the one lost with the killed animals, then such killing is permissible. For this to be true and, as I have argued, faithful to Utilitarianism, 'lead to' must relate the killing with its consequences. However, killing animals, *by itself*, does not "lead to the creation of other animals which …" Some additional actions are required, like making animals reproduce, taking good care of the newborn, and so on. Thus, the plausible sense in which killing animals leads to such and such is by being one action among a sequence of actions which has that consequence. Yet, in this sense, (R) tells us that the value of the new animal's life is not a consequence of the killing (nor of the other isolated actions). According to consequentialism, then, the new animal's life is irrelevant to the normative status of the killing. Therefore, (1) is false, for although the killing leads (in the specified sense) to the valuable state of affairs, this has no bearing on its permissibility.[6]

An obvious question now arises: why should a consequentialist accept (R)? Well, if one cares about stopping the RA, then this already counts in its favour. But of course that this alone will seem rather *ad hoc*. Moreover, without any further support, (R) is also too strong a claim just to deal with a problem for utilitarians concerned with the ethics of killing animals. Nevertheless, I believe that we can say more in favour of (R).

III.b   Two Further Motivations

Firstly, without a restriction like (R) the consequentialist allows the value of conditional actions to trivially influence the value of actions on which they are conditional. And I think that this is untenable. Consider an example of *value sabotage*. You did an intuitively permissible action like saving a person's life. Now suppose that someone killed another person on the condition that your saving was successful. Then your saving may not be permissible after all, for its normative status depends on the overall value of those two actions. It is odd that the value of an action depends not just on the things that it brings about, but also on the things *chosen* to be brought about by it (*mutatis mutandis* for *value improvement*, where the conditional action allegedly improves the condition action). To be clear, in these

---

[6] Were the argument stated with the consequence relation, (R) implies that the antecedent of (1) is false, making premiss (2) false. Interestingly, Persson (2017, 78-9) agrees that raising good lives cannot compensate killing good lives, "for while the latter could be done by means of a single act, the former cannot." But Persson leaves unclear why performing various acts cannot compensate a single one. My proposal is a step to explain this.

cases, conditionalising is itself a result of agency. Therefore, contrarily to non-agential conditional events, the conditional action can occur without the condition action.

This way of influencing the normative status of actions is too trivial to be acceptable. Even if every non-agential consequences of one's action would be good/bad, conditional actions could always overturn the balance. So, unless one is prepared to abandon a view of agency as being tightly connected to individual responsibility, consequences of other actions should not be treated like an action's non-agential consequences.

Consequentialists, then, can make a relevant distinction between consequences of sequences of actions and consequence-related events tracing back to a single action.[7] (R) does just this by preventing that consequences of sequences of actions bear on the normative status of the isolated actions. Still, (R) does not depart from the basic idea that, to evaluate an action, consequences are all that matter. In this way, although (R) restricts the received view about what counts as consequences of an action, we remain on consequentialist ground.[8]

Secondly, (R) has other useful applications for consequentialists. I will point out two.[9] The first regards the so called "accordion effect" of action. In brief, the worry is that the same set of events can be appropriately described in various ways that are such that the action in one description contains some of its consequences in another description. Adapting an example from Miller (1987), consider these two descriptions of what Jones did:

(a) Jones tells a lie.

(b) Jones saves a life.

If (a) and (b) are correct descriptions of Jones' action, then, assuming that the relevant value is in (b), consequentialists can only account for the normative status of the action via description (a); in contrast, non-consequentialists will care if (b) follows (or not) some rule. If

---

[7] E.g. pushing a person on the street is not permissible because someone decides to benefit that person if you push her; however, it would be if, say, by pushing her, a bullet happens to miss her. In the latter case, but not in the former, the valuable consequence is a result of your action alone. Note also that all I said is compatible with both single or multiple agent sequences of actions.

[8] Smart (1956) distinguished "extreme" and "restricted" Utilitarianism by, respectively, having a focus on single actions or on classes of actions. My suggestion is similar, but I am distinguishing single actions from sequences thereof and disregarding their being subsumed under a rule. Thus, in Smart's sense, Utilitarianism with (R) is still extreme.

[9] The aim here is just to motivate (R)'s acceptance beyond the RA and not to exhaust its usefulness. But I also envisage other applications, e.g. to a more commonsensical consequentialist account of blameworthiness.

this is sound, then, as Oldenquist (1966, 183) puts the problem, "whether we appeal to rules or to consequences to determine the rightness or wrongness of a particular action is of no moral significance." That is, the notion of consequence is left without distinctive normative relevance.

Given (R), however, the accordion can only be stretched so much: although we can agree that, say,

> (*) Jones deceives the intending murderer

is also a correct description of what he did, we cannot say the same of description (b). The reason is that (b), but not (*), forces us to recognise multiple actions – whether or not a life is saved also depends on the intending murderer's action. Thus, (R) prevents the accordion from stretching beyond descriptions involving single actions.

Finally, another useful application of (R) pertains the cluelessness objection (Lenman 2000). In brief, the objection is the following: since the consequences of our actions are normally spread in time and space in a way that surpasses our knowledge, then we have no clue about what we ought to do. What seems to be a perfectly permissible action, like sparing the life of a pregnant woman, might actually be impermissible because such action happens to have the consequence of not preventing the birth of a future terrible dictator and all his atrocities.

Again, with (R) at disposal, the consequentialist has a line of response: consequences of the dictator's actions are not consequences of sparing his ancestor. We remain clueless about the consequences of sequences of actions that contain our actions as parts. Yet, given that such consequences are not consequences of our actions alone, we are not required to know them (we could not). And since they have no bearing on the normative status of our actions, ignoring them does not imply that we are in the dark about what we alone ought to do.[10]

---

[10] There is a reply if we can be clueless even if there are no sequences of actions involved. But the burden of proof is with those who think that single actions can have massive causal ramifications and that most of our actions are like that. But note that the claim here is modest: if (R) can mitigate this problem (as well as the accordion effect), then its acceptability goes beyond its stopping of the RA.

IV   Objections

One tempting objection to my way of blocking the RA is that, somehow, we can automate the sequence of actions that together lead to the new animal's valuable life. In this way, it seems that there would be a single action, e.g. the press of a button, that leads to the killing of one animal and to the raising of another satisfying the relevant conditions. Thus, the value of what would otherwise be a sequence of actions is, in the automation case, the value of a single action. Since this, apparently, would not involve a sequence of actions, (R) would not apply and, therefore, it seems that the killing would be permissible (given that the press of the button would).[11]

This objection fails because it overlooks one crucial action (or sequence), namely, setting up the automation, making it seem that (R) would not apply when in fact it does. Hence, the valuable state of affairs would still be a consequence of a sequence of actions.

Perhaps one serious objection is that the RA can be restated in a way that bypasses (R). One might say that it does not matter whether or not the valuable state of affairs is a consequence of the killing, for as long as the whole sequence brings about such state of affairs, then, replacing an animal, that is, the whole sequence, is permissible. In other words, we shift the evaluation focus from actions to sequences of actions. And since I do not deny that the valuable state of affairs is a consequence of the sequence, then it seems that I have to agree that it determines (or contributes to) the normative status of the sequence.

But is this an objection to my proposal? The goal was to argue that, contrarily to widespread agreement, standard act-Utilitarianism does not imply the RA. After all, this was the target of those who used the RA against Utilitarianism (e.g. Pluhar 1982; Regan 2004). To achieve that goal I proposed a novel way, via (R), to stop the RA. But I did not claim that every utilitarian view with (R) stops the RA. It may well be the case that a global utilitarian view, that is, one which allows every sort of thing as evaluative focus, implies the RA. At the very least, the objector has to argue that a utilitarian should accept sequences of actions as evaluative focus. This comes with difficulties.

The said shift of evaluative focus requires completing and making sense of the new, reformulated principle:

---

[11] I had thought of this objection before, but I thank Melinda Roberts for mentioning it to me and thereby confirming my intuition that it was something I had to address.

(C*) A sequence of actions is permissible iff it brings about more value than any other alternative _____ available to the agent.

The natural move is to fill the blank with 'sequence', but do agents have alternative sequences of actions to choose from? Maybe just in single agent sequences, for an agent cannot choose a sequence that involves other people's actions (otherwise he would know how others would act).[12] And while single agent sequences are enough to formulate the RA (but seriously limiting its application), we would still need a systematic account of the normative relation between sequences and the actions composing them. Without such account, that the consequences of a sequence are good overall is not enough for its permissibility, since it may be the case that a single impermissible act stains the sequence of which it is part.

V   Conclusion

The value of the new animal's life should be a consequence of killing another animal if the RA is to be successful. Yet, I argued that such value, given the restriction on consequences I presented, is not a consequence of the killing. Therefore, the first premiss of the RA is false. Since that restriction is quite strong and, apparently, *ad hoc*, I offered two distinct motivations for it: one axiological and one of usefulness. I then considered and replied two plausible objections, the last of which hints at further work on coordinated actions and on the normative relation between sequences and the actions composing them.

I should conclude by stressing that even though I could not assess here the relative merits of others ways to disassociate Utilitarianism and replaceability, my proposal does not give up of any of the usual suspects like those other ways do – hedonism, maximisation or the total view. And while I am sure that other objections might be raised, I think that this utilitarian proposal against the RA is worthy of being discussed in more detail.[13]

---

[12] What about coordinated actions? Here seems possible to choose a sequence involving other people's actions because everyone agreed to act in such and such manner and so the agent seems reasonably informed in a way that does not preclude the sequence from being an alternative action. I have no answer to this.

[13] Meanwhile, following comments from Theron Pummer, Bruno Jacinto, José Mestre and Pedro Galvão, to all of whom I am thankful, I became aware of other difficulties, and also possible developments, of the view presented here. I hope I can address them in the future.

Acknowledgement

References

[1] Bratman, Michael. 2006. "What is the Accordion Effect?" *The Journal of Ethics* 10 (1-2): 5-19. DOI: 10.1007/s10892-005-4589-3.

[2] Chappell, Richard. 2015. "Value Receptacles." *Noûs* 49 (2): 322-32. DOI: 10.1111/nous.12023.

[3] Delon, Nicolas. 2016. "The Replaceability Argument in the Ethics of Animal Husbandry." In *Encyclopedia of Food and Agricultural Ethics*, edited by Paul Thompson, and David Kaplan. Dordrecht: Springer. DOI: 10.1007/978-94-007-6167-4_512-1.

[4] Hart, H. L. A., and A. M. Honoré. 1967. *Causation in the Law*. Oxford: Oxford University Press.

[5] Lenman, James. 2000. "Consequentialism and Cluelessness." *Philosophy & Public Affairs* 29 (4): 342-70. DOI: 10.1111/j.1088-4963.2000.00342.x.

[6] Miguel, Ricardo. 2016. "What is the problem of replaceability?" In *Food futures: ethics, science and culture*, edited by I. Anna S. Olsson, Sofia M. Araújo, and M. Fátima Vieira, 52-8. Wageningen: Wageningen Academic Publishers. DOI: 10.3920/978-90-8686-834-6_6.

[7] Miller, Arthur. 1987. "Acts and Consequences: Squeezing the Accordion." *Metaphilosophy* 18 (3-4): 200-7. DOI: 10.1111/j.1467-9973.1987.tb00853.x.

[8] Oldenquist, Andrew. 1966. "Rules and Consequences." *Mind* 75: 180-92. DOI: 10.1093/mind/LXXV.298.180.

[9] Persson, Ingmar. 2017. *Inclusive Ethics: Extending Beneficence and Egalitarian Justice*. Oxford: Oxford University Press.

[10] Pluhar, Evelyn. 1982. "On Replaceability." *Ethics and Animals* 3 (4): 96-105.

[11] Regan, Tom. 1983/2004. *The Case for Animal Rights*. Berkeley: University of California Press.

[12] Singer, Peter. 1979/2011. *Practical Ethics*. 3rd edition. Cambridge: Cambridge University Press.

[13] Smart, J. J. C. 1956. "Extreme and Restricted Utilitarianism." *The Philosophical Quarterly* 6 (25): 344-54.

[14] Višak, Tatjana. 2013. *Killing happy animals: explorations in utilitarian ethics*. Basingstoke/Hampshire: Palgrave Macmillan.

[15] _____. 2016. "Do Utilitarians Need to Accept the Replaceability Argument?" In *The Ethics of Killing Animals*, edited by Idem, and Robert Garner, 117-35. Oxford: Oxford University Press.