

# The Perils of Earnest Consequentializing

SERGIO TENENBAUM

*University of Toronto*

Portmore argues in Portmore (2011) that we can consequentialize any ethical theory; according to Portmore, any remotely plausible view is deontically equivalent to a consequentialist view. Let us assume for a moment that this is true. There are at least two ways to see the significance of this result. I'll call the first interpretation the 'dismissive consequentializer interpretation'. On this view, the the possibility of consequentializing any view shows that the debate between deontology and consequentialism is misconceived; there is no interesting difference between consequentialism and nonconsequentialism.

On the other hand, the 'earnest consequentializer' takes these results as a vindication of consequentialism; in particular, these results supposedly show that we can have the intuitive appeal of consequentialism without having to bite bullets. Earnest consequentializers claim they can get all they want: theoretical satisfaction without carving organs for the greater overall good or letting one's children suffer for the sake of a slight overall gain for strangers. Portmore is clearly on the side of the earnest consequentializer, and this is the view I wish to examine.

Let us start by looking at three ways in which one can challenge the earnest consequentializer:

1. Failure of Extensional Equivalence

The consequentializing project fails if it cannot deliver the same deontic statements as some plausible nonconsequentialist theory.

2. Triviality

The consequentializing project fails if the consequentialized versions are just notational variants of, or in some significant way derivative from, the nonconsequentialist theories.

### 3. Explanatory Inadequacy or Inversion

One might accept the result and yet think that nonconsequentialist views are the right ones because either (i) the consequentialist view fails to explain why an action is permissible or obligatory (explanatory inadequacy); or, (ii) the truth of consequentialized formula is explained by the nonconsequentialist theory rather than the other way around (explanatory inversion).

These charges form some kind of trilemma for the earnest consequentializer. You could make your view substantive enough so that it avoids (2) and (3), but then the substantive view is likely to have unwanted consequences, and fall prey to (1). One can instead broaden one's notion of outcome, stretch one's conception of what is valuable,<sup>1</sup> relativize the good in various ways, etc. until we get consequentialism to deliver the verdicts we want. But this might land us into (2) or (3). We might have no understanding of what is good from the point of view of the agent, at a time, except that it is an outcome that results from the agent's acting permissibly. And even if we do have independent access to the resulting notion of the good, it seems more plausible that something is good relativized to the agent and to a time because this is among the actions he is permitted to perform rather than the other way around. In this paper, I argue that Portmore's view cannot escape this trilemma. I have no illusions that I establish this point conclusively; it will take much more than I can do here to address all of Portmore's powerful arguments.

Let us examine a definition of (maximizing) act-consequentialism that Portmore provides at a certain point. Even if this is not Portmore's final version of consequentialism, it'll help us frame a few issues:

[ACT] An act is morally permissible if and only if its outcome is not outranked by that of any other act alternative (86).

Now this definition tells us nothing about how 'outcome' or 'outranked' are understood, so it is not clear that it escapes the triviality objection; perhaps 'outrank' needs to be defined in terms of the permissibility of different acts. It might seem obvious that we understand outcomes being ranked in a way that has nothing to do with moral permissibility; in fact, we can think that this follows from our understanding of a partial ordering, in which outcomes are ranked from lower to higher. However, Portmore denies that 'x outranks y' is transitive for at least his favourite version of the 'outranking'

---

<sup>1</sup> Portmore's version of consequentialism does not rank outcomes in terms of their value. But it's simpler to formulate the dilemma in these more familiar terms.

relation (148). So at least in non-maximizing versions of consequentialism, we cannot rely on a partial ordering for our understanding of outranking. In fact, we get very few constraints on this relation. The only one I can find is that it is asymmetric (89). So it is not clear that there any constraints on the ‘outranking’ relation other than the constraint that permissible acts are such that their outcomes are not outranked by any other. Moreover if [ACT] is to allow the range of views that Portmore wants to accept, then ‘outranking’ might be relativized to anything; in fact, there is no principled restriction of the “relativizing operators”. Portmore suggests two: agent and time. But given this understanding of [ACT], it is unclear why relativization should be restricted to those. In fact, Portmore effectively commits himself to relativizing ranks to possible worlds. In explaining how consequentialism can allow for the possibility of dilemmas Portmore says that:

To do so, the consequentialist must hold that whether  $O_1$  outranks  $O_2$  or vice versa depends on whether I perform  $A_1$  or  $A_2$ . More specifically, the consequentialist must hold that  $O_1$  outranks  $O_2$  if and only if I perform  $A_2$  and that  $O_2$  outranks  $O_1$  if and only if I perform  $A_1$  (90-1).

However, since ‘outranks’ is asymmetric, and ‘ $O_1$ ’ and ‘ $O_2$ ’ refer to broad outcomes (that is the possible worlds that would be actualized by my choosing the respective acts), it cannot be that they outrank each other unless we relativize ‘outrank’ to something. It obviously cannot be agent or time. The only difference in the two cases is which possible world gets actualized, so the only way we can deliver this result is to relativize ‘outranks’ to possible worlds. Allowing act-relative consequentialism into the consequentializing project significantly weakens it.

None of this establishes that we have a trivial definition of ‘outcome X outranks outcome Y’. But it does undermine some of the intuitive appeal of the earnest consequentializer; some of the most appealing consequentialist ideas are left behind. Agent-relativity had already taken away the advantage that consequentialism seemed to be less ‘self-indulgent’ or ‘self-centered’ than its nonconsequentialist counterparts. This version of consequentialism can no longer claim that it always enjoins the agent to pursue the best outcome. Moreover, the hopes for explanatory priority dim at least at this general level. Portmore explains the intuitive appeal of his version of consequentialism as follows:

If our actions are the means by which we affect the way the world goes, and if our intentional actions necessarily aim at making the world go a certain way, then it is only natural to suppose that what we have most reason to do is determined by which way we have most reason to want the world to go (56).

But I am not sure why this is intuitive. It is simply false if the act is not part of the outcome, if the consequences of my act must be causally downstream from my act. When I go running I aim at nothing causally downstream from my action. Of course, very few, if any, modern consequentialists have such a restrictive (though natural) view of consequences. Portmore says that “The aim need not be anything having to do with the causal consequences of the act. The aim could be nothing more than to bring it about that one performs the act.” (56). But is it intuitive that when I go running for its own sake, there is a proper substitution instance of “go in a certain way” that validates the above formula?<sup>2</sup> With a bit of ingenuity, we could find a way to characterize an aim as an instance of ‘making the world go a certain way’ that is satisfied if and only if my aim to run is satisfied. But even if there is nothing wrong with stretching our understanding of consequences to accommodate the possibility of actions done for their own sake, we cannot claim intuitive appeal for the resulting apparatus. It is perhaps intuitive that all actions are chosen for their own sake or for their consequences (now understood in a more ordinary way). But this thought does not favour consequentialism over nonconsequentialism.

Even if the general intuitive appeal of consequentialism is lost, Portmore’s specific version might turn out to have a compelling combination of initial plausibility, extensional adequacy, and explanatory power. So let us turn to Portmore version. I’ll focus here not on Portmore’s Dual Ranking Act Consequentialism but on the Teleological Conception of Reasons that underlies it. Here is TCR:

(TCR) S has more reason to perform  $a_i$  than to perform  $a_j$  if and only if, and because, S has more reason to desire that  $o_i$  obtains than to desire that  $o_j$  obtains (56).

(TCR) has the advantage of not depending on a notion of agent-relative goodness that has been the focus of criticism lately (most notably by Schroeder (2007)). Outcomes are ranked not by the ‘better than’ relation, but by the ‘more reason to desire’ relation. This is a very important contribution. It takes the debate over consequentializing into a new direction as Portmore’s version of consequentialism need not say that the right action maximizes the good. Portmore (2011) puts forward an ingenious account of how a consequentializer can eschew ‘desirability’, ‘good’, and their cognates entirely.

But since in his account ‘more reason to desire’ does the heavy lifting, we should look at this relation more closely. There are two related

---

<sup>2</sup> For doubts about an affirmative answer, see Lavin and Boyle (2010).

suspensions that connect to the trilemma above: 'more reason to desire that an outcome obtains (or that a possible world is actualized)' is not clearly an ordinary notion and thus we might suspect that our understanding of this notion is simply derivative of our understanding of a reason to act, or as Portmore puts it 'more reason to perform (an action)'. On the other hand, one might suspect that to the extent that there is a notion of 'more reason to desire' that is not derivative, it does not validate (TCR).

In a common understanding of 'desire', desire is not an 'all-out' attitude, an attitude that does not allow for contradictory contents. I can desire that **p** and that **not p** without irrationality. It is not so clear how to assess the strength of reasons to desire in this sense of 'desire'. It is natural to think that, say, my reason to X is weightier than my reason to Y just in case it is true that, *ceteris paribus* if I am in a situation such that doing X and doing Y is incompatible, I should do X rather than Y. But since desiring X and desiring Y are never incompatible, we cannot use this kind of criterion to measure the strength of the reason. Although I might have to forego doing X because my reasons to do Y are stronger, I need never stop desiring X because my reasons for desiring Y are stronger. One could respond that the weight of my reason to desire is determined by the weight of the corresponding reason to act, but this would give us the opposite explanatory direction from (TCR).

In some cases, it seems that we can determine the weight of a reason to desire by the degree to which someone distances herself from ideal reason responsiveness by failing to have the desire in question. So, for instance, a parent typically has reason to desire the welfare of her children. So a parent who has no desire to help her children is not responding to a reason she has even if, in the current circumstances, she cannot, or should not, bring about any improvement to her child's welfare. Someone who failed to have this desire would exhibit a serious failure of reason-responsiveness.

But can we say the same thing about the desire to give jobs to those who are most qualified for the job? Is there a reason to desire such a thing whose weight is explanatorily prior to the reasons I have to engage in such behaviour? Even if there is such a reason, the failure to have a desire for the well-being of one's children would be a greater, or at least not a lesser, failure of reason-responsiveness. Yet, it could easily happen that I must choose to give a job to the most qualified person even if it would greatly promote the welfare of my child if I gave it to her. Perhaps one can argue that in this specific situation I have more reason to desire giving the job to the most qualified person. But this response raises some doubts that 'more reason to desire' is the more basic explanatory notion; it seems that what would explain that I have more reasons to desire giving the job to the most qualified person is my reasons to perform this action (or the greater value of this alternative) rather than the other way around.

Do I have a reason to desire that, say, a possible world in which Larry gets a birthday present and Mary has her toe stubbed be actualized? I desire that Larry get a birthday present and I do not desire that Mary have her toe stubbed. But we are supposed to be asking here if I desire, all things considered, that a certain possible world be actualized and an answer requires that I know much more about this world. Even if we could imagine a possible world in which all that happens is that Larry gets a birthday present and Mary has her toe stubbed, it isn't clear that we should have an attitude such as desire towards it. My overwhelming temptation in being asked whether I have a reason to desire that such a world be actualized is to ask "Compared to what?". It is unclear that I have reasons to desire all things considered this outcome, let alone that we can attach weights to these reasons that I can use to compare with other outcomes.

Postulating a *comparative* desiderative attitude ranging over outcomes is relatively unproblematic. However, such an attitude is better described as 'desire more', or better, given that it is a rather different attitude from 'desire', we can use its more common name 'prefer'. In fact, in explaining his view Portmore freely moves from 'more reason to desire' to 'reason to prefer', so it might seem like just a stylistic accident that he uses 'more reason to desire' as opposed to 'reason to desire more', or 'reason to prefer'. But I think Portmore is completely right in choosing 'reason to desire' over 'reason to prefer' in the official formulation of the view. It is unlikely that what we ought to prefer is explanatorily basic. The reasons to prefer an outcome very broadly conceived to another are likely to be a function of some other reasons or evaluative facts. Replacing 'reason to desire' with 'reason to prefer' in the (TCR) formula does not settle the question of explanatory direction. But if 'A has reason to prefer X over Y' is explained in terms of the reasons to perform X (and the reasons to perform Y), then (TCR) cannot be true given the obvious explanatory circularity it would entail. If 'A has a reason to prefer X over Y' is explained in terms of some evaluative or desirability facts (for instance, the facts that constitute the fact that 'X is better than Y'), then we lose the gains that Portmore had made by providing a statement of consequentialism that does not rely on such notions. So for (TCR) to be true and have the theoretical advantages that Portmore proposes for it, it really needs to be stated in terms of 'more reasons to desire', and not 'reasons to desire more'.

I have been suggesting that the reasons here do not always add up in the right direction; the general reasons to desire the welfare of my child seem not to be less weighty than the general reasons to desire that the most deserving person receives the job. A related, more general problem can be put this way: Portmore wants to rely on a notion of 'reason to desire' that is agent-relative and that is explained neither by reasons to perform actions nor by agent-relative values (or facts about desirability). It is then natural

that Portmore rejects a conception of desire that identifies a desire with a disposition to act; after all, it would be rather implausible to think that reasons to act are explained by reasons to be disposed to act rather than the other way around. But if desires involve further features, there is no guarantee that such features will not generate a different normative profile. So, for instance, frustrated and satisfied desires might have conceptual ties to certain emotions, and the connected emotions might add weight to having a desire for certain outcomes even when there are more reasons to promote other outcomes. This might be the case with my reason to desire the welfare of my children. For instance, perhaps some negative emotions when a desire fails to be satisfied are constitutive of a desire for X. Perhaps there are weightier reasons to have these emotions when I fail to bring about the welfare of my children than when I fail to give the job to the most qualified applicant. In sum, we cannot simply equate ‘more reason to desire’ with ‘reason to desire more’. The former seems the right notion for the explanatory ambitions of earnest consequentializing, but it is the latter that can hold promise for delivering on extensional equivalence.

Portmore argues that a view in which explanatory priority always goes in the same direction is preferable to a view in which it does not, and that (TCR) gets at least some instances right:

Suppose that a5 is the act of putting my money into a savings account that yields 2% annually and that a6 is the act of putting my money into a savings account that yields 1% annually... In this case ... what explains the fact that I have more reason to perform a5 than to perform a6 is the fact that I have more reason to desire that o5 obtains than to desire that o6 obtains... Surely no one would argue that the reason I have for preferring the outcome in which I am prudentially better off is that this is the outcome in which I have done what was prudent (78).

This example does seem, at least at first, to favour the TCR direction over the ‘anti-TCR’ one. However, in other cases the explanatory priority seems to go in the opposite direction. We could start with the classic case in which I can only save five people by killing a different person. Here it seems that it is reason to refrain from killing that explains the reasons to desire the outcome in which I do not kill. The idea that the explanatory priority in this case runs the TCR direction is no more plausible. Moreover, reasons to earn money are instrumental reasons and thus explained by other reasons; I have reasons to dance, to walk the streets of Paris, to ride a horse, etc. All these things cost money and thus the reasons to engage in these actions explain my reasons to earn more money. Once we get to non-instrumental reasons, there’ll be much less bullet biting on the left to right direction of explanation. In sum, such examples can at best leave the TCR and the anti-TCR direction on a par.

Furthermore, and more importantly, the argument seems to rest on an imprecise characterization of the opponent's view. Note that the anti-TCR direction would not entail that "the reason I have for preferring the outcome ... is that this is the outcome in which I have done what was prudent"; it at most entails that the fact that I have reason to prefer this outcome is explained by the fact that I have more reason to perform this action. The opponent's view does not imply anything about *what the reason for wanting the outcome is*. Arguably, what explains that I have a reason to X or that I have a reason to desire X are the facts that constitute this reason. My reason to save money is that money is needed to go to Paris. My reason to go to Paris is that it'll be an enriching experience; my reason to help is that my friend is in need, etc. Portmore's opponent needs only to say that these facts explain that I have a reason to X without the intermediacy of a reason to desire.

### References

- Lavin, D. and Boyle, M. (2010). Goodness and Desire. In Tenenbaum, S, editor, *Desire, Practical Reason, and the Good*, pages 158–199. Oxford University Press, New York.
- Portmore, D. W. (2011). *Commonsense Consequentialism: wherein morality meets rationality*. Oxford University Press, New York.
- Schroeder, M. (2007). Teleology, Agent-Relative Value, and 'Good'. *Ethics*, 117(2):265–295.