

UC Berkeley

ISUS-X, Tenth Conference of the International Society for Utilitarian Studies

Title

“Can Consequentialization Advance the Cause of Consequentialism?”

Permalink

<https://escholarship.org/uc/item/8vc0r9ct>

Author

Sachs, Ben

Publication Date

2008-08-21

“Can Consequentialization Advance the Cause of Consequentialism?”

Benjamin Sachs

National Institutes of Health

Abstract

Recent work on consequentialism has revealed that theory to be more flexible than previously thought. Consequentialists have shown how their theory can accommodate intuitions about the value of other things besides happiness, the agent-centeredness of some constraints, and the importance of how value is distributed. This is usually thought to be of great benefit for consequentialism. I want to cast doubt on this assumption. I begin by putting forward the strongest possible statement of consequentialism’s flexibility—the claim that, *whatever* set of intuitions our best non-consequentialist theory accommodates, we will be able to construct a consequentialist theory that can do the same while still retaining the virtues of consequentialism. (This process is called consequentialization.) I then argue that if this is true then most likely the non-consequentialist theory with which we started will have the virtues of consequentialism. So while consequentialization makes consequentialism more appealing, it makes non-consequentialism more appealing too.

I. Introduction

Several authors have argued that the case-specific judgments generated by any non-consequentialist normative ethical theory can be delivered in a consequentialist framework.

Some of them believe that from the possibility of “consequentializing” non-consequentialist theories it follows that we need to rethink either the usefulness or the very existence of the distinction between consequentialism and non-consequentialism.¹ Douglas Portmore, on the other hand, argues that what follows is that there is a version of consequentialism with a particular multifaceted appeal that, until now, it was thought normative ethical theories could not achieve.² Supposing there is, however, has the cause of consequentialism been advanced? I will demonstrate here that whatever the particular sort of appeal Portmore or others manage to show, through the process of consequentialization, that consequentialism possesses, there is a good chance that that same process of consequentialization will show that there is a non-consequentialist theory that is appealing *in precisely the same way*. Therefore, consequentialization is, for consequentialism, a double-edged sword.

II. How to Consequentialize

Portmore’s version of the consequentialization program is an effort to demonstrate the truth of the following claim:

For any remotely plausible non-consequentialist theory, *M*, there is a consequentialist counterpart theory, *M**, that yields, in every possible world, the exact same set of deontic

¹ James Dreier, “Structures of Normative Theories,” *Monist* 76(1): 19-30; Jennie Louise, “Relativity of Value and the Consequentialist Umbrella,” *Philosophical Quarterly* 54(217): 518-536; Michael Smith, “Neutral and Relative Value after Moore,” *Ethics* 113(3): 576-598.

² Douglas Portmore, “Consequentializing Moral Theories,” *Pacific Philosophical Quarterly* 88 (2007): 39-73.

verdicts that *M* does, including not only such verdicts as “permissible” and “impermissible,” but also such verdicts as “supererogatory.”³

The truth of this claim, the Deontic Equivalence Thesis (DET), would demonstrate that one can build into act-consequentialism certain features with which it has standardly been thought inconsistent, including options, agent-centered constraints, time-centered constraints and supererogation. In this section I want to explain Portmore’s method for constructing an act-consequentialist theory that can accommodate agent-centered constraints. Having shown how this is done, I will leave it to the reader to deduce how Portmore’s theory could handle the other features just listed.

For an agent-centered constraint to exist is for it to be the case that there is a certain action that one may not perform even if doing so would prevent multiple other tokens of that same type of action from being performed. On a standard understanding of what act-consequentialism is, it appears to not be able to accommodate agent-centered constraints. Act-consequentialists believe that an action is permissible for an agent, *A*, if and only if there is no other action available to *A* that would have a better outcome. Given this claim, no matter how bad an action’s outcome is it will be permissible (obligatory, in fact) to perform it if doing so will prevent that same action from being undertaken multiple times. After all, if one person’s *X*-ing leads to a bad outcome, then five people’s *X*-ing must lead to a *really* bad outcome.

³ “Consequentializing,” p. 54, original italics removed. Portmore makes it clear elsewhere that by “consequentialist” he means act-consequentialist.

Portmore proposes an alternative understanding of act-consequentialism on which an action is permissible for an agent, A, if and only if there is no other action available to A that would be *better-relative-to-A*. And what it means for an action to be better-relative-to-A is that the object-given reasons A has to prefer the outcome that would result from it are better than the object-given reasons A has to prefer the outcome that would result from any other available action. On this version of act-consequentialism, the ordering of outcomes is relativized to the agent. Because of this, it might be impermissible for an agent to perform an action with a bad outcome even if performing the action would prevent five other people from doing the same. For instance, suppose A faces the choice of whether to kill an innocent person in order to prevent five other agents from each killing an innocent person. We might reasonably argue that A has better object-given reasons to prefer that he not engage in killing, since doing so would cause him to lose his sense of integrity. If the argument goes through, then A is required to refrain from killing. In other words, A is under an agent-centered constraint to not kill.

Let this suffice as an argument for the compatibility of a certain form of act-consequentialism—agent-relative act-consequentialism—with agent-centered constraints. I will assume that this argument is sound. Furthermore, let's suppose that agent-relative act-consequentialism can be shown to be compatible with the other kinds of deontic verdicts with which act-consequentialism has long been thought incompatible, such as options, time-centered constraints, and supererogation, and that therefore DET is true. We need to make these assumptions so that we can proceed on to the question we set out to answer: Supposing DET is true, what follows?

One thing that would follow from the truth of DET is that act-consequentialism would be more appealing. It is widely accepted that there is something extremely enticing about any act-

consequentialist theory—something compelling idea in it that even its detractors detect.

Portmore believes the compelling idea, which he calls “the permissibility of maximizing view” (PMV), is this:

It is always permissible (in the knowledge-supposing sense) for an agent to act so as to bring about the highest-ranked available outcome, i.e. the outcome that she has better object-given reasons to prefer about all other available alternatives.⁴

The possibility of consequentialization (i.e., the truth of DET) shows how, contrary to what most theorists have always thought, we can deliver any plausible set of deontic verdicts in an act-consequentialist framework, thereby allowing us to hold on to PMV. Thus, agent-relative act-consequentialism can give us in a single package the compelling idea—if Portmore is right about what the compelling idea is—and whatever set of deontic verdicts we find most intuitively acceptable.

One might wonder, as Mark Schroeder has, whether on closer inspection Portmore’s argument succeeds in establishing this conclusion.⁵ Again, it is not my intent here to critically examine the argument, so I will have to push this question to the side. What I want to ask is this: Supposing Portmore’s argument goes through, what has been accomplished? I will argue now that if there is a version of act-consequentialism that accommodates PMV and some set of deontic verdicts, then there is a version of non-act-consequentialism that does the same.

⁴ Portmore, “Consequentializing,” p. 50.

⁵ Mark Schroeder, “Teleology, Agent-Relative Value, and ‘Good’,” *Ethics* 117 (2007): 265-295.

III. What Do We Get out of Portmore's Argument?

Suppose a non-act-consequentialist theory, N, can be consequentialized, meaning that there is an act-consequentialist theory, C, that delivers N's deontic verdicts while asserting PMV. In most cases, for the non-act-consequentialist to show that N has the same virtues as C will be as easy as helping herself to the fruits of the Portmore's labor. First, let's give the name " Δ " to the set of deontic verdicts yielded by N. In order to consequentialize N, Portmore, as we have seen, puts forward a theory of outcome-ranking that demonstrates that Δ is consistent with PMV. Δ and N are consistent with each other, of course, so by the transitivity of consistency, N is consistent with PMV. Therefore, N, like C, is consistent with both Δ and PMV. What this reveals is that all the non-act-consequentialist has to do in order to show that her theory has the same two virtues as its consequentialist counterpart is to accept Portmore's theory of outcome-ranking and let logic do the rest.

This strategy will be available except when the non-act-consequentialist cannot take on Portmore's theory of outcome-ranking. This condition will hold only when N itself is committed to claims about which outcomes are better than others, as occurs in rule-consequentialism and some versions of non-consequentialism such as W.D. Ross's theory.

Where does this leave us? Portmore, we are assuming, has shown how any non-act-consequentialist theory can be consequentialized. We have shown that some of those non-act-consequentialist theories—i.e., the ones that don't claims about the ranking of outcomes—embody exactly the same two virtues as their consequentialist counterpart: consistency with Δ

and PMV. This conclusion is interesting, but there is a stronger, more important conclusion just a one-premise argument away.

The premise is: for any set of deontic verdicts, Δ , Δ can be delivered by a non-consequentialist theory, N , that is compatible with PMV. There will usually be a number of ways to create such an N , but at least one method will always be available: the construction of a ten commandments-style list of DOs and DON'Ts, one for every member of Δ . For instance, if “it is impermissible to kill Jones” is a member of Δ , just put “DON'T kill Jones” on the list of rules in N . Since commandment-style rules make no claims about which outcomes are better than others, N will be compatible with PMV. From this premise we may infer the conclusion that if there is an act-consequentialist theory that can deliver our preferred set of deontic verdicts while remaining compatible with PMV, then there is a non-act-consequentialist theory that can do the same. (We can safely assert this conditional because we know its consequent to be true; its consequent is an instance of the universal claim that was the premise.) Thus, far from promoting the cause of act-consequentialism, Portmore-style consequentialization actually puts act-consequentialism and non-act-consequentialism—at least the best versions of each—on equal footing. This is what is meant by the idea that consequentialization is a double-edged sword.

The question I now want to ask is whether we can say the same of consequentialization more generally. If we dispense with whatever aspects of Portmore-style consequentialization are peculiar to Portmore, and focus instead on only the necessary elements of a consequentialization program, will we find that consequentialization per se cannot advance the cause of consequentialism? From here forward our concern will not be Portmore's concern about whether consequentialization can advance the cause of act-consequentialism in particular. We

will instead employ the familiar consequentialism/non-consequentialism distinction and inquire as to whether consequentialization can advance the cause of the consequentialism, full stop.

IV. What Do We Get Out of Consequentialization Per Se?

First we should get clear as to what we mean by “consequentialization per se.” As I said earlier, some theorists push the view that consequentialization prevents us from drawing a useful consequentialism/non-consequentialism distinction. For obvious reasons we need not concern ourselves with this use of consequentialization. We need focus only on the versions of consequentialization that purport to yield a previously-undiscovered, attractive version of consequentialism.

Next we need to identify the necessary elements of a consequentialization program in this sense. We can start with the basic observation that to consequentialize is to take the most plausible set of deontic verdicts and put forward a value theory⁶ that shows that these verdicts can be delivered in a consequentialist framework. If this can be done, it will constitute the discovery of the most attractive version of consequentialism yet. Such a consequentialist theory, however, would not necessarily have an edge over the best non-consequentialist theory. To get that advantage, the consequentializer has to show that the new consequentialist theory is compatible with some compelling idea with which all versions of non-consequentialism that deliver that same set of deontic verdicts are incompatible. That is to say, she has to show that the two theories constitute a counterexample to the following universal claim (UC):

⁶ Construed broadly enough such that a theory of outcome-ranking counts as a value theory.

For any set of deontic verdicts, Δ , and any compelling idea, if there is a consequentialist theory that can deliver Δ while remaining compatible with the compelling idea, then there is a non-consequentialist theory, N , that can do the same.

Our final task, then, is to determine if she will be able to find such a counterexample. Whether this can be done depends, of course, on what the compelling idea is. Since our concern here is consequentialization per se, we cannot presume that we know what the compelling idea is; different variants of consequentialization might put forward differing compelling ideas. What, then, are the possibilities?

As a first pass we can say that consequentialism's compelling idea either is or is not a proposition that is part of normative ethical theories themselves. If it is not, then the consequentializer is going to have to say that it is a metaethical claim. The reason for this is that the consequentializer needs the compelling idea to be a proposition with which some normative ethical theories (including typical versions of consequentialism) are compatible and others (including the best versions of non-consequentialism) are not. If the compelling idea is not itself a proposition about right and wrong, good and bad—the sort of proposition out of which normative ethical theories are built—then it is going to have to be a proposition about metaphysics, epistemology or normativity that has some bearing on which propositions about right and wrong and good and bad we can accept. Otherwise, it will be irrelevant to the choice between consequentialism and non-consequentialism. Any metaphysical, epistemological or

normative claim that is relevant to the identification of the correct normative ethical theory is, I shall stipulate, a metaethical claim.

I am going to set aside the possibility that the compelling idea might be a metaethical claim, for the simple reason that the class of metaethical claims is just too large to be surveyed in the way we need. Given the great variety of metaethical claims that are out there, I cannot imagine what general considerations one might adduce to show that the compelling idea simply cannot be a metaethical claim. We will have to wait until the consequentializers identify the best candidates. It bears mentioning, however, that it seems unlikely that the compelling idea will be a metaethical claim. Two reasons for this. First, some metaethical claims have no bearing on what normative ethical theories we can accept, and the ones that do usually do not make the cut between acceptable and unacceptable at the line between consequentialism and non-consequentialism. Second, one would think that whatever draws philosophers to consequentialism would be an element of consequentialism itself; after all, how many of us have really taken the time to consider which metaethical claims are consistent with consequentialism? And for those of us who have bothered to do this, isn't this because we were *already* attracted to consequentialism? If so, then consequentialism's compelling idea should be found in consequentialism itself.

The more likely possibility, and hence the one more worthy of our attention for now, is that the compelling idea is a claim found in normative ethical theories. My goal for the rest of this paper is to make a prima facie case against this possibility. My method will be simple: I will examine one-by-one the various elements of normative ethical theories—the various kinds of claim that normative ethical theories sometimes make. The object of each examination will be to

determine whether, if the compelling idea were that kind of claim, there might be a counterexample to UC. My goal is to make a prima facie case for the truth of UC.

First, however, we need to identify the elements of normative ethical theories. I am going to be generous here, allowing in some kinds of claim that one might doubt are really parts of normative ethical theories themselves. (Given my argumentative strategy, this generosity can only help my opponent.) I am going to allow that normative ethical theories have four elements: First, one or more moral principles. Second, evaluative claims. Moral principles, combined with non-moral facts and/or evaluative claims, yield the third part of a normative ethical theory: a set of deontic verdicts. The fourth part is a determination claim—a claim about what kinds of fact determine the truth of the deontic verdicts. Determination claims are to be distinguished from moral principles in the following way: whereas moral principles state necessary and/or sufficient conditions for an action's rightness or wrongness, determination claims identify the facts that *ground* an action's rightness or wrongness.

Having identified the elements of normative ethical theories, we can begin our search for counterexamples to UC.

1. Moral principles

Suppose we construct a non-consequentialist theory, N, to fit Δ . (We will always be able to do this, as I explained earlier.) Is there a chance that it will be incompatible with the compelling idea, on the assumption that the compelling idea is a moral principle? We already know from Section III that we can construct N such that its moral principles are compatible with PMV, which itself is a consequentialist moral principle, and a quintessential one at that. Generalizing

inductively, it seems quite probable that if the compelling idea is any consequentialist moral principle at all then we'll be able to construct N such that its moral principles are compatible with it. Therefore, in the case in which the compelling idea is a consequentialist moral principle, and thus the antecedent of UC is true, the consequent of UC will probably be true.

2. Evaluative claims

Again, suppose we construct a non-consequentialist theory, N, to fit Δ . As we have seen, this can always be done, and furthermore can always be done such that N makes no evaluative claims. And so it is trivially true that, if the compelling idea is an evaluative claim, we can construct N such that it delivers Δ and is compatible with the compelling idea. Hence, we can be certain that if the compelling idea is an evaluative claim, then UC's consequent, and thus UC itself, is true.

3. Deontic verdicts

Recall that for any set of deontic verdicts, Δ , we can construct a non-consequentialist theory, N, to deliver Δ . So if the compelling idea is a deontic verdict, then it is trivially true that we can construct a non-consequentialist theory that will deliver Δ and be consistent with the compelling idea. Consequently, if the compelling idea is a deontic verdict, then UC's consequent is true, and so is UC.

4. Determination claims

I argued in Section III that for any set of deontic verdicts, Δ , if there is a consequentialist theory that can deliver Δ while remaining compatible with PMV, then there is a non-consequentialist theory, N, that can do the same. One might wonder whether N deserves the label “non-consequentialist” if it doesn’t deny a clearly consequentialist moral principle like PMV. Furthermore, I suggested earlier in this section that no matter what consequentialist moral principle the compelling idea might be, it is difficult to imagine that we won’t be able to construct N such that the compelling idea is compatible with it. Thus I am relying heavily on a heretofore implicit claim that non-consequentialist theories can accept consequentialist moral principles. My view, which I take to be intuitively plausible enough not to require defense, is that a moral theory that accepts a mixture of consequentialist and non-consequentialist moral principles is non-consequentialist if its determination claim is non-consequentialist. A determination claim is non-consequentialist just in case it holds that the facts that determine the permissibility of an action are something beside the goodness of the state of affairs that the action would bring about.⁷

Imagine, for instance, a theory that includes the following two moral principles:

- a) An act is wrong if its performance under the circumstances would be disallowed by any system of rules for the general regulation of behavior which no one could reasonably reject as a basis for informed, unforced general agreement.
- b) An act is wrong if there was an alternative to it available to the agent that would have brought about a better state of affairs.

⁷ We should of course construe “state of affairs” broadly enough to encompass facts about the act itself. That way a determination claim remains consequentialist even if it takes into account, for instance, the badness of a right being violated.

Such a theory is non-consequentialist if it also holds:

- c) The permissibility of an action is determined by whether its performance under the circumstances would be disallowed by any system of rules for the general regulation of behavior which no one could reasonably reject as a basis for informed, unforced general agreement.

Being lenient about moral principles requires being strict about determination claims. To defend UC while insisting that non-consequentialist theories can accept consequentialist moral principles, we are going to have to say that for a theory to be non-consequentialist it must contain a non-consequentialist determination claim. And this raises the prospect that if consequentialism's compelling idea is a determination claim then we will not be able to construct N such that it is compatible with that claim. This prospect is made all the more likely by the fact that determination claims are by their very nature exclusive. To illustrate the point, consider the following determination claim:

- d) The permissibility of an action is determined by whether there was an alternative to it available to the agent that would have brought about a better state of affairs.

We have to say that c and d are incompatible unless we want to admit the possibility of moral overdetermination (a move that wouldn't help us in the end, since a theory that accepted both c and d probably wouldn't deserve the label "non-consequentialist"). By contrast, there is nothing troublesome about saying that a and b, the two moral principles, are compatible.

So if consequentialism's compelling idea is a determination claim, then we have our counterexample to UC. But I doubt that consequentialism's compelling idea is a determination claim. On the contrary, in fact, consequentialism's determination claims are usually seen as one of its main drawbacks. If Jones stalks and kills an innocent person in cold blood, consequentialism tells us that it's the resulting state of affairs that determines whether Jones's action was impermissible. Thus, even if consequentialism gives the right answer, it will seem to give it for the wrong reason.

This is a well-worn complaint, of course, but strikingly few theorists have made an effort to respond to it. The literature contains just a handful of avowals of consequentialist determination claims, and even fewer defenses of them. Given this, it is hard to imagine that any determination claim can take much of the credit for the longstanding, ongoing popularity of consequentialism, and therefore constitute consequentialism's compelling idea.

Now one might object that I have unfairly stacked the deck against consequentialism's determination claim by presenting an insufficiently general version of it. I said just now that it seems that consequentialism will give the right answer for the wrong reason if it says that Jones's cold-blooded killing is impermissible on the grounds that Jones could have brought about a better state of affairs by doing something else. While such a determination claim looks less than compelling, we might think that the general claim of which it is an instance is unimpeachable. That claim would be something along the lines of "morality is about doing what's best." One might even suppose that what accounts for the literature's dearth of defenses of consequentialism's determination claim is that the claim, when abstracted to a sufficient level of generality, is just obviously true.

This appearance, I submit, is deceptive. Notice, first, that “morality is about doing what’s best” is not a determination claim. A determination claim identifies some facts as determining the truth of some proposition(s). “Morality is about doing what’s best” doesn’t do this. The determination claim that best captures the spirit of this proclamation is, “the fact as to whether X-ing will bring about the best available state of affairs determines whether X-ing is permissible.” Already this begins to look more debatable.

And things gets worse for the consequentialist, for her determination claim, as we are now interpreting it, is ambiguous in a troubling way. There is an ambiguity in the term “best”—the question being whether we are talking about moral or non-moral bestness. Suppose it’s the former. In this case, the consequentialist determination claim would read: the fact as to whether X-ing will bring about the morally best state of affairs determines whether X-ing is permissible. But a consequentialist cannot say this, for the following reason. If she believes in moral bestness, then she believes in moral goodness. And if anything is morally good, surely a permissible act is, *qua* permissible act. Or, at the very least, permissible acts are morally better than impermissible ones. Therefore, the consequentialists need to tell us which acts are permissible *before* she can tell us which things are morally good or bad.

Therefore the consequentialist must instruct us to interpret “best” in her determination claim as signifying “non-morally best.” So her determination claim is now, “the fact as to whether X-ing will bring about the non-morally best state of affairs determines whether X-ing is permissible.” I very much doubt that many philosophers have been drawn to consequentialism by such a claim.

Thus, whether we look at consequentialism's determination claim from afar or in an applied context, it doesn't look like the sort of thing that deserves the "compelling idea" label. Therefore, we can admit that there would be a counterexample to UC if the compelling idea were a determination claim, and yet be reasonably confident that there is no counterexample to UC.

I have argued in this section in support of the claim, which I labeled UC, that for any set of deontic verdicts, Δ , and for any compelling idea, if there is a consequentialist theory that can deliver Δ while remaining compatible with the compelling idea, then there is a non-consequentialist theory, N, that can do the same. That we can construct an N so as to make it compatible with Δ I take for granted, having argued in Section III that we will be able to do this no matter what Δ is. So we have focused our efforts on determining just what shape the compelling idea might take such that non-consequentialist theories cannot accommodate it. We have found dispositive reasons to believe that N will not be incompatible with the compelling idea if the latter is an evaluative claim or deontic verdict. And we put forward an inductive argument that gives us a compelling reason, though not a dispositive one, to doubt whether N could be incompatible with the compelling idea if the compelling idea were a moral principle. Finally, we found a strong reason to think that the compelling idea is not a determination claim, meaning that we needn't worry about which determination claims N is or is not compatible with. Thus, we have a prima facie case for the truth of UC. The burden of proof is now on the consequentializer to undermine this case by showing one of the following three things: 1) The compelling idea is a consequentialist moral principle with which non-consequentialism per se is incompatible, 2) the compelling idea is consequentialism's determination claim, or 3) the

compelling idea is a metaethical claim with which consequentialism is compatible and non-consequentialism per se is incompatible.

IV. Conclusion

Consequentialism has long been subject to criticism for failing to accommodate widespread intuitions about deontic verdicts. Post-Bentham consequentialists such as Mill, Moore, Sen and Scheffler have aimed to show that consequentialism has more flexibility than it has been credited for. Carrying on in this tradition, Douglas Portmore has argued that act-consequentialism is flexible enough to accommodate widespread intuitions on deontic verdicts, and furthermore it can do so while remaining compatible with the idea that, according to Portmore, makes act-consequentialism compelling: that it is always permissible to maximize the good (PMV). I have argued that this flexibility, if act-consequentialism does indeed have it (a question I have left to the side), would turn out to be a double-edged sword. While it would indeed allow act-consequentialism to accommodate widespread intuitions on deontic verdicts while remaining compatible with PMV, it would show that the element of non-act-consequentialism that was always thought to make it incompatible with PMV—its deontic verdicts—does not do so. And so non-act-consequentialism would also be able to accommodate widespread intuitions on deontic verdicts while remaining compatible with PMV.

I have argued, further, that even though we aren't certain what consequentialism's compelling idea is—Portmore might be wrong after all—we have prima facie justification for believing that if we can construct a version of consequentialism that delivers our preferred set of deontic verdicts while remaining compatible with the compelling idea, whatever it is, then we

can construct a version of non-consequentialism that does the same. Our justification arises from the fact that, so long as consequentialists affirm a set of deontic verdicts with which non-consequentialists are comfortable, non-consequentialists needn't deny any consequentialist claims, including most of their candidate compelling ideas. The only exception to this is that non-consequentialists do have to deny the consequentialists' determination claim. However, that claim is probably not consequentialism's compelling idea.

The cash value of all this is that most likely we can construct a version of non-consequentialism that has whatever virtues consequentializers claim for their spruced-up versions of consequentialism. And so consequentialization apparently does not advance the cause of consequentialism.