Consequentializing Group Membership:  
A Reply to Causal Impotence

Word count (excluding notes and bibliography): 2,984

Abstract: In this paper, I defend a new way for consequentialists to respond to cases of collective harm. First, I explain why some of the popular proposals, such as the “threshold response,” fail to provide fully general reasons in these cases. I then suggest how consequentialism could potentially resolve the issue by including group membership as a morally relevant consequence. According to my proposal, becoming a member of a group that collectively causes more harm than good is a bad consequence of an action. Thus, we always have a pro tanto reason to abstain from contributing to a collective harm. I argue that this can explain what goes wrong when someone claims to be blameless on grounds of causal impotence, and I maintain that this proposal coheres with common intuitions about collective action without rejecting consequentialist commitments.

Keywords: Collective action, causal impotence, environmental ethics, applied ethics, consequentialism

“Even if an act harms no one, this act may be wrong because it is one of a set of acts that together harm other people.”¹

1. Introduction

There has been a great deal of recent discussion dedicated to questions about collective harm and what an individual ought to do in such cases.² It has proven somewhat difficult for consequentialists to explain why someone ought to perform a particular act when she can plausibly claim that her action will not make a difference. For instance, what reason do I have to reduce my greenhouse gas emissions when I have little reason to believe that my individual actions will have an impact on the collectively caused harms of climate change?³

A variety of responses to this problem have been proposed, but most of them have limitations that make them fall short of providing reasons that apply to all cases. Perhaps the most popular reply is the idea that individuals have reason to believe that their actions could make a difference because the outcome in question will be triggered when a critical threshold is

¹ Derek Parfit, *Reasons and Persons*, 70. My position is the result of an attempt to make sense of this conclusion from Parfit’s “Five Mistakes of Moral Mathematics” — the *locus classicus* for the problem of collective action.

² It would be impossible to provide an exhaustive list of references. For some recent discussions of collective action and causal impotence, see Gunnemyr 2019, Dietz 2016, Chignell 2015, Nefsky 2012, and Kagan 2011.

³ Greenhouse gas is one popular topic for the problem of collective harm. See, for example, Gunnemyr 2019, Kingston and Sinnott-Armstrong 2018, and Sandberg 2011. Of course, there are many other examples of collective harms and causal impotence. The consumption of intensively farmed animals is another widely discussed case.
crossed. It is then suggested that such cases can be handled by ordinary considerations of expected utility. Even if the probability of triggering the outcome is quite small, the harm might be extremely severe, and individuals therefore have a reason to avoid actions that might bring about disastrous outcomes. Although this response seems quite promising at first glance, it cannot adequately handle all instances of collective harm. First, it has been argued that such a proposal speaks only to a proper subset of cases; it might not apply to ones involving imperceptible harm (such as Parfit’s harmless torturers). Second, even in cases where the expected utility argument does apply, the chance of making a difference is sometimes so small that it can effectively be ignored.

If the current responses to the problem are limited in the ways I believe they are, then there are some worrisome cases (perhaps many of them) where these replies will fail to give individuals any reason to act. In this paper, I defend a new proposal according to which consequentialists always have a pro tanto reason to avoid performing actions that contribute to collective harms. The response I propose involves an argumentative tactic that is called “consequentializing.” The idea behind this move is that consequentialists can alter an ethical prescription by broadening the set of consequences that are considered to be morally relevant. By building more things into the theory of the good, consequentialists can theoretically win over their opponents by accommodating intuitions about putative counterexamples.

To use this tactic as a solution to the problem of collective harm, I suggest that we consequentialize group membership. More specifically, when an agent performs an action that makes her a member of a group that collectively causes more harm than good, we should consider that to be a bad consequence of her action. In order to make sense of this proposal, I will address some of the issues it raises. For starters, there will be some groups, which I call “mere aggregates,” that do not count as groups in the relevant sense. I will have to explain what justifies that distinction in order to avoid counterintuitive results. There will also be some difficult questions about how to evaluate the amount of harm done by joining a group and how to rank outcomes, but I believe that such questions can, in principle, be answered. I make some suggestions on how this could be done.

---

4 There are many proponents of the threshold argument. For a representative sample, see Kagan 2011, Norcross 2004, and Singer 1980.
The structure of my paper is as follows. In section 2, I explore some of the replies to the problem of causal impotence that have been proposed, and I explain why they cannot always provide reasons for action in these cases. In section 3, I show how my view has the advantage of applying to all cases of collective harm. I conclude by replying to some objections to my proposal and suggesting how to address the obstacles it might face.

2. Some Popular but Unsatisfying Solutions

Although a wide array of solutions to the problem of collective harm have been suggested, I will limit myself to two of them here. I have narrowed the scope in this way for two reasons. First, I am limiting my attention to consequentialist proposals, since my aim here is to see whether or not consequentialism can provide reasons for action in these cases. Second, I have chosen consequentialist proposals that have been widely discussed and are thought by many to be the most promising replies.

I will begin with the expected utility argument. This is arguably the most popular proposal that has been defended by consequentialists. Shelly Kagan, Peter Singer, Alastair Norcross, and many others have argued that considerations of expected utility give us a compelling reason to abstain from eating factory farmed meat. Most versions of the argument share the following assumptions. First, the wrongness of the action depends on the collectively caused harm (e.g., the animal suffering, environmental impacts, etc.). Second, these harms will not be mitigated by a single person becoming a vegetarian (when that action is considered in isolation), but a sufficiently large drop in the demand for factory farmed meat will result in a significant reduction of the harms (and vice versa: a sufficiently large increase in demand will cause more animals to be bred and slaughtered). If the above claims are true, then there is a chance that my order of a chicken sandwich will cross the threshold and cause tens of thousands of chickens to suffer greatly. The marginal utility of the chicken sandwich over the vegetarian alternative is surely outweighed by the suffering of the chickens, and I therefore have a reason to abstain from eating it. Ordering the chicken sandwich has negative expected utility.

---

7 There have been some attempts to deal with the problem by making use of Kantian universalizability tests. See, for example, Chignell 2015 and Shafer-Landau 1994. Shafer-Landau argues that the Kantian test does not offer any reason to abstain from eating factory farmed meat; he suggests that a virtue-based proposal might fare better.

The argument has some intuitive plausibility, but it faces several objections. The first problem is that it does not apply to all cases of collective action. Julia Nefsky has argued that there are many examples of collective harm that do not fit the above description of the factory farming case. Although the expected utility reply might provide reasons in instances of this kind, which she calls “triggering cases,” there are some cases of “imperceptible differences” to which the solution does not apply at all. Nefsky shows how Kagan’s argument ultimately relies on the mistaken assumption that all cases of imperceptible differences are actually “triggering cases in disguise.”\(^9\) She argues convincingly that Parfit’s “harmless torturers” and “drops of water” pose serious problems on this count. In such cases, it might really be true that an individual’s actions do not make a difference of the sort that Kagan has in mind.

The second line of objection maintains that even in cases where the threshold argument does apply, the chance of making a difference is sometimes so small that it ought to be ignored. This is notoriously problematic for the case of moral vegetarianism (and personal greenhouse gas emissions). The factory farming example has long served as a paradigm instance where the threshold reply is supposed to offer consequentialists a reason to avoid contributing to the collective harm. But the pushback on this case has been particularly strong. Given the vast size of the industry and its apparent insensitivity to changes in demand, the numbers required for change are very large. And as the number gets bigger, the odds of my purchase making a difference get smaller. If we are told in response that we ought to take small chances seriously (as Parfit, Norcross, Singer, and Kagan tell us we should), then the proposal might prove too much. The result of taking small chances seriously will almost certainly be overly demanding. Above and beyond vegetarianism, we might be risking the threshold chicken simply by eating at a restaurant. They could be cooking the vegetarian entrée in chicken stock, and that might have been the threshold chicken!\(^10\)

It should also be noted that, when it comes to the expected utility proposal, the details matter. There will be cases where we can be reasonably certain that our action will not trigger the bad outcome, and thus there is surely greater expected utility for eating the meat in such cases. For instance, we might know enough about how a certain restaurant operates and when they plan to place an order for the next crate of chickens. There could be wastes, buffers, and inefficiencies

---

\(^9\) Nefsky, 365.

\(^10\) Chignell 2015 makes this point, and I think it is certainly right. Taking small chances extremely seriously would require us to abide by an unlivable standard.
that cause some purchases to be totally ignored.\textsuperscript{11} This would mean that one of the underlying assumptions is false. It was wrong to suppose that every purchase moves the needle of aggregate demand; some purchases are completely invisible and make no difference even on the underlying dimension.

Another reply to causal impotence focuses on groups as the relevant agents and thereby moves the moral question from the individual to the group level.\textsuperscript{12} After all, if the harm in question was caused collectively (and individuals are causally impotent) then we might suspect that it is groups who deserve our attention. The downside of this solution is that it might not be able to offer individuals any reason to behave one way rather than another. Indeed, some versions of this proposal bite the bullet and completely absolve individuals of any wrongdoing by placing blame solely on group agents.\textsuperscript{13} This would mean that the group proposal gives no answer to the question of why I, as an individual, ought to reduce my greenhouse gas emissions, abstain from eating factory farmed meat, etc. It would tell groups how they ought to act, but those reasons might not trickle down to individuals.

If this is right, then neither of the above proposals are able to offer reasons for individuals to act in all cases of collective harm. The group agent proposal locates some wrongdoing and offers prescriptions, but they do not necessarily have force at the individual level. The threshold argument does not apply to all cases, and there are skeptical worries even in the cases to which it does apply. My aim in this paper is to offer a new proposal that covers all cases and always generates a reason for an individual to act in such a way that she does not contribute to a collective harm.

3. Consequentializing Group Membership

The basic thrust of my proposal is quite simple. Consequentialists have a long tradition of expanding their theory of the good in order to accommodate things that are excluded from narrower theories (such as hedonistic ones). In some cases, this can be done as a way of producing a consequentialist theory that is extensionally equivalent to an alternative theory. For

\textsuperscript{11} See Budolfson 2019.
\textsuperscript{12} See, for example, Kiloren and Williams 2013.
\textsuperscript{13} Ibid. Kiloren and Williams explicitly state that the group agent acts wrongly, but none of the individuals in the group violate any moral obligation. According to other proposals, such as Alexander Dietz’s, groups have obligations, and some of these obligations do trickle down to individuals but only if others are doing their part. See Dietz 2016.
instance, consequentialists could include “treating a person as a mere means” as a harmful consequence and give it considerable weight in order to accommodate recalcitrant intuitions that lead some people to prefer Kantian theories.\(^{14}\)

Broadly construed, outcomes are defined as states of affairs, and it is up to the consequentialist to decide on the criteria for determining why one state of affairs is preferable to another. Monistic theories of value, such as hedonism and preference-satisfaction, are very well known (and well liked), but pluralistic theories have enjoyed some popularity as well. They sometimes include things such as pleasure, knowledge, beauty, autonomy, etc. For my purposes here, it does not matter which theory of the good we start with, but we will end up with a pluralistic one after combining it with the consequence that I want to include: group membership.

Some actions have the consequence that it makes the agent a member of a group. There are times when this is quite obvious (e.g., when I pay dues for the APA), and there are times when it is more subtle (e.g., when I buy a chicken sandwich and renew my membership in the group of people who purchase factory farmed meat).\(^{15}\) What I want to suggest is this: one good consequence of an action is that it makes the agent a member of a group that causes more good than harm. Conversely, it is a bad consequence of an action if it makes the agent a member of a group that causes more harm than good. The implications of this for the problem of collective harm should be quite obvious. In every case of collective action, there is now a pro tanto reason to avoid contributing to a collective harm. Whether it is Parfit’s harmless torturers, purchasers of factory farmed meat, or emitters of greenhouse gases, we have a reason not to join such groups. The causal structure of the case no longer matters; whether it involves triggering a critical threshold or an imperceptible harm, we have a reason to opt out of the action if it will make us a member of a harmful group.\(^{16}\)

I think this suggestion has strong intuitive plausibility. For instance, if there is a large group of schoolyard bullies that is taunting and beating someone, Sally has a reason not to join

---

\(^{14}\) This tactic, when taken to the extreme, eventually leads to questions about whether or not there are limits to what kinds of things can be consequentialized and concerns about vacuity. See Brown, “Consequentialize This.” It seems clear to me that those concerns are overblown. The tactic be very useful, and it is more than a mere “gimmick,” as Nozick once called it. Nozick, *Anarchy, State, Utopia*. Cf. Vallentyne, “Gimmicky Representations of Moral Theories” and Davies, “Structures of Normative Theories.”

\(^{15}\) It might seem controversial whether or not this is a group at all, but I explain my reasons for this below.

\(^{16}\) Of course, conversely, we would have a pro tanto reason to join a beneficial group such as the one described in Parfit’s “drops of water” case.
the mob even if she is extremely certain that her action of taunting from a distance will not make any difference to the first-order harm (however that is defined: pain, unsatisfied preferences, etc.). Of course, consequentialists who are skeptical of my proposal might be eager to find other reasons for condemning Sally. They could claim that Sally risked causing additional suffering to the victim (giving her action negative expected utility). Or they could point out that there are future harms at stake. She might not be able to comfort the victim afterwards; she might bully other children in the future, and so on. But even if these harms are completely absent from the case, Sally’s action still seems wrong. And my proposal would help us explain why. Our condemnation of Sally’s behavior is not contingent on the amount of suffering that she caused (or could have reasonably expected to have caused). We should condemn her even if we are quite certain that she did not increase suffering (or expected suffering) in the slightest.\footnote{We could also modify the harmless torturer case in order to motivate the intuition. The more torturers there are, the worse the world is, even if we hold constant the amount of suffering that the torture victim experiences.}

4. Objections and Replies

One immediate worry concerns the definition of a group. It would be a serious problem for my proposal if membership in a set of agents who perform some action type was, in and of itself, sufficient for moral consideration. For instance, let’s say that someone detonates a bomb killing hundreds of innocent people. That might make it the case that people who push buttons today cause more harm than good. But that should not give me any reason to abstain from pressing a button on the elevator. I similarly doubt we have a pro tanto reason to abstain from shooting a gun even though it is quite plausible that that the set of people who shoot guns today will cause more harm than good.

But there is a principled way of excluding such sets of agents from our moral consideration. Unlike the APA, the Senate, and greenhouse gas emitters, the group of people who push buttons today do not collectively cause anything. The distinguishing feature of morally relevant groups is that they collectively bring about some effect (by means of a joint effort, whether coordinated or not). This is what separates a bona fide group like “consumers of factory farmed meat” from a mere aggregate such as “people who shoot guns today.” Even though it is likely that the latter causes more harm than good, it is not a group in the sense I am suggesting because the consequences of their actions, though very harmful, were not caused
collectively. Global warming, health care policies, and APA meetings, however, are indeed the products of collective action. This also gives us a reason to refine the principle I am proposing. One formulation goes as follows: we have a pro tanto reason to abstain from joining a group that causes more harm than good. But it is important to clearly state that we have such reasons from abstaining only if the group collectively causes more harm than good. The group does not count (a) if it does nothing collectively or (b) if the members of the group cause more harm than good but they do so on an individual basis.

The last objection might be the most serious one. One requirement of a consequentialist theory is that it must be able to rank outcomes in order to make judgments about which actions to prescribe. My suggestion about group membership has specified a state of affairs that could be considered good or bad, but I have not said anything about the kind of weight that should be assigned to this consequence. My first reply to this problem is to point out that other theories of the good face this problem as well. It is incumbent on both pluralistic and monistic theories to ask some difficult questions in cases where goods are in competition. I propose using two criteria by which to judge the consequence of group membership. First, the greater the harm that the group collectively causes, the worse it is to join that group. Second, it is, ceteris paribus, a weightier consequence to join a small group than it is to join a large one. The reasons behind both of these ranking criteria should be quite obvious. The smaller the group, the larger our contribution to the collective action. And the more harm or benefit the group causes, the more seriously we must consider avoiding or joining the group.

6. Conclusion

This proposal gives us a new tool for thinking about problems of collective action. It might not yield decisive considerations in all cases, but it does, at the very least, speak to every case of collective action, regardless of its causal structure. For those who think that the problem of collective harm is a serious one for consequentialists (as many do), this suggestion could help move the debate forward by providing a new ground for action in such cases.

---

18 Even hedonistic theories face hard questions of this kind. But preference-satisfaction theories do as well. This problem is a familiar one, and it should not trouble us too much.
Bibliography


