

Participation and Superfluity

Jan Willem Wieland

Rutger van Oeveren

Why act when the effects of one's act are negligible? For example, why boycott sweatshop or animal products if doing so makes no difference for the better? According to recent proposals, one may still have a reason to boycott in order to avoid complicity or participation in harm. Julia Nefsky has argued that accounts of this kind suffer from the co-called "superfluity problem", basically the question of why agents can be said to participate in harm if they make no difference to it. This paper develops and responds to Nefsky's challenge.

Keywords: participation, membership, complicity, collective action, superfluity

In Journal of Moral Philosophy

1. Preamble

Some people boycott sweatshop clothes or animal products because they think that by doing so they will make a difference for a better world. Others don't share that conviction, and don't boycott. Alternatively, they *do* boycott, though not because they believe they will make a difference—they believe they won't—but for some other reason. Specifically, they boycott because they *do not want to be complicit*. These agents are motivated by the idea that by boycotting sweatshop clothes they won't be complicit in exploitation, or that by boycotting animal products they won't be complicit in animal suffering.¹

The question is: does such a reason justify their conduct, and does it render alternative conduct unjustified? This question afflicts not just cases of boycotting, but more generally cases of collective action. In all such cases, individuals might not make a difference, but they might still be complicit, or, as we will understand this notion, participate in a group of agents who *together* make a difference. Julia Nefsky has argued that such participation-based reasons are undermined by, what she calls, the "superfluity problem." The contribution of our paper is two-fold: first, we spell out the superfluity problem in greater detail and show how it poses a serious worry (§§2-3), and, second, we develop and defend our solution to it (§§4-6).

As we take it, individuals do have reasons to act in certain ways in collective contexts because of participation-based considerations (possibly among other considerations). You have a reason to act, roughly, because by acting you avoid participation in harm. If you buy sweatshop and animal products, you are part of the group that sustains exploitation and animal suffering. You are one of them. You are not one of them, and avoid membership of such groups, in contrast, if you boycott. We take this to be a strong enough reason to boycott. Even so, in this paper we won't provide an ultimate justification for such participation-based reasons. Our very ambition, instead, is to show that these reasons are *not* undermined by Nefsky's superfluity problem.

¹ Cf. DeGrazia (2009), Martin (2015), Driver (2015), Harman (2015), Budolfson (2015), McPherson (2015), among others. It is also one way of cashing out Singer's claim that we should boycott animal products in order to make our lives "as free of cruelty as we can" (1975: ch. 4).

One important thing before beginning. In this paper, we will talk about pro tanto reasons for action, rather than their wrongness (following Nefsky 2017: 2744-5). Still, if one has a reason to boycott, and there are no stronger reasons against boycotting, then it might well be wrong not to boycott. Moreover, we will assume that such reasons can have a forward-looking function (“do I have a reason to boycott?”) as well as backward-looking one (“did I have a reason to boycott?”).

2. Superfluity problem

Consider the following streamlined cases:²

Drops of Water. 1000 wounded men are lying out in the desert, suffering from intense thirst. An equal number of altruists have a pint of water. They could pour these pints into a water-cart, which will then be driven to the desert, and the water will be shared equally among the wounded men. Each altruist who donates a pint enables each wounded man to drink only slightly more water, a drop. Since one such drop will not benefit even a very thirsty man, no altruist decides to donate her pint.

Harmless Bandits. In a village, 1000 unarmed tribesmen are about to have their lunch. An equal number of hungry armed bandits descend on the village. They tie up the tribesmen, then have a look at their lunch. Each bowl of food contains 1000 baked beans. Instead of each bandit eating a single plateful, each takes just one bean from each plate. They leave after eating all the beans, pleased to have done no harm.

In Harmless Bandits many people together *harm* others, while in Drops of Water many people together *fail to benefit* others.³ We will need an account

² Edited from Parfit (1984: 76) and Glover (1975: 174-5) respectively.

³ The consumer case may be taken in both ways. Many consumers together may harm the workers abroad by demanding products made in bad circumstances,

for both kinds of cases, and the problem is this. In Drops of Water, we may think that (1) the altruist has a reason to add her pint only if this benefits the men in the desert, and that (2) adding her pint does not benefit these men (more), and yet that (3) the altruist still has a reason to add her pint. Similarly, in Harmless Bandits we may think that (1) the bandit has a reason to refrain from stealing a bean from a plate only if stealing it harms the tribesmen, and that (2) stealing a bean does not harm the tribesmen (more), and yet that (3) the bandit has a reason to refrain from stealing a bean.

If, in each case, (3) holds, then, by standard logic, (1) or (2) has to go. But which? Many participants in the debate concede (2). In collective actions involving numerous agents, the relevant outcomes are insensitive to single individuals. Whether the tribesmen will be harmed, or the wounded men be helped, will depend on what *many* people decide to do, not on what you individually do. Nor does the extent to which they are harmed or benefited depend on what you individually do. Whether or not you donate or steal, the wounded men or the tribesmen will be harmed or helped all the same.⁴

Another strategy may be to deny (1), and to maintain that reasons for action do not require difference-making. In that case, the altruist still has a reason to add her pint even if that makes no difference to the situation of the men in the desert, or the bandit still has a reason to refrain from stealing a bean even if that makes no difference to the situation of the men in the village. This strategy might seem promising (at least to non-act consequentialists), though developing it is more problematic than it may seem. In particular, accounts that deny (1) face the superfluity problem. As Nefsky writes:

or they may fail to benefit the workers by not demanding products made in better working conditions.

⁴ Does one not even *marginally* harm or benefit? One can do so only if harms and benefits can be “imperceptible”, and that is controversial, see Parfit (1984: ch. 3) and the subsequent debate. We will simply accept (2) for the purposes of this paper.

The problem is that [such accounts] do not seem to work as long as your act would be instrumentally⁵ superfluous. And yet that seems to be exactly what we are accepting when we accept that your act cannot make any difference. (2015: 264)

Importantly, the problem does not simply consist in the reply “there is no reason to act so long as it makes no difference.” That would merely beg the question against accounts that deny the importance of difference-making. Then what *is* the problem? Nefsky does not provide one general statement of it, and in the following we will spell out her suggestions in *two* different ways—or, if you like, there are simply two separate problems. To explain this, let us consider the fairness account that Nefsky (2015: §§5-6) discusses:

Fairness. S has a reason to do X if by doing X, S is doing her fair part.⁶

For example, by adding her pint, the altruist is doing her fair part, and by refraining from stealing the bean, the bandit is doing her fair part as well. It would be unfair to take the beans or to keep one’s pint while others do not. As a proponent of this account clarifies: “It is a matter of pulling my weight in what we all ought to be doing.” (Cullity 2000: 17)

Does the altruist pull her weight if she donates her pint? Here is the first problem for the fairness account.⁷ The proponent of this account agrees that donating her pint is superfluous. It is not that she has a reason to donate because by doing so she will help the wounded men. It is that she has a reason to donate because by doing so, she is pulling her weight. What does this mean? Arguably, something along the following lines: that one contributes a reasonable amount to further the collective endeavor of helping

⁵ An act is “instrumentally” superfluous when it is superfluous with respect to realizing a relevant outcome. We will define the notion of “superfluity” in §2.

⁶ Throughout the paper, ‘S’ stands for an individual agent, and ‘X’ for an action or omission (donating one’s pint, refraining from stealing a bean).

⁷ In this paper, we will focus on the superfluity problem, and leave aside further potential problems with the accounts. E.g. it is not unfair to keep one’s pint if *everyone* does so (cf. Cullity 2000: 18). If so, the fairness account generates a reason to donate only *after* others start to donate (cf. Nefsky 2015: 255).

the wounded men. But whatever “reasonable” might mean, if acting in the relevant way won’t make any difference, then it is unclear that it contributes to the collective endeavor at all. Thus, given the superfluity of X, it is unclear that I am pulling any weight, and consequently that, so understood, I am doing my fair part (cf. Nefsky 2015: 257).

One might suggest that fairness does not amount to “pulling one’s weight”, but rather to “making things equal”. By donating one’s pint when others have done so as well, one makes things equal.

However, this definition generates a second problem for the fairness account. After all, there are different ways to make things equal that do not involve donating one’s own pint. As Nefsky imagines: “As far as bare equality in the resources of the original pintholders is concerned, it would be just as effective if I pour my pint onto the ground.” (2015: 259) The problem here is that we don’t want an account that says we have a reason to pour our pint onto the ground. We want an account that says we have a reason *to donate our pint*.

These two problems constitute a serious challenge to the fairness account, and potentially to any alternative account of the following format: S has a reason to do X, even when X makes no difference, if by doing X, S is F. Let us state these problems more generally in terms of two types of counterexamples, which we will label respectively the “Undercutting Counterexample” and the “Alternatives Counterexample”:

Undercutting Counterexample. Doing X is superfluous. Hence: S does not count as “F”. Hence: the account fails to generate a reason to do X. Yet, S has a reason to do X, which still has to be explained.

Alternatives Counterexample. S can also do X* rather than superfluous X. If S does X*, S also counts as “F”. Hence: the account fails to generate a reason to do X, i.e. rather than X*. Yet, S has a reason to do X, which still has to be explained.

Counterexamples of both types imply that the given account is *insufficient* for certain reasons for action. But each type does so in a different way. In the Undercutting Counterexample, doing X does not imply being F, and so the account gives no reason to do X (i.e. not even a pro tanto reason). In the

Alternatives Counterexample, being F does not imply doing X (for there are different ways to be F), and so the account gives no reason to do X (not even a pro tanto one) if one can do something else that makes one F as well.

Hence, instead of talking about “the” superfluity problem (as Nefsky does), we think there are in fact two distinct problems. The schemas just given will help in raising them. Before showing how further accounts are undermined by them, we would like to clarify three things.

First point: why think that accounts should give us a reason to do X (e.g. to donate one’s pint rather than to pour it on the ground)? Nefsky writes:

We need an explanation of why each of us specifically has reason to do an act that is of the sort that if enough of us do it, this will collectively cause (or, depending on the case, prevent) the morally significant outcome of concern. Finding a reason to add a pint to the cart that is just as much a reason to instead wear a T-shirt that expresses one’s support or solidarity does not do this. (2015: 263)

Hence, the idea is that accounts should give us a reason to do X, because if enough people do X, then there will be enough people to realize the good outcome (or prevent the bad one).

We will accept this with one qualification. We think that sometimes it is fine if individuals fail to do X, but rather do something else that makes one F, namely in cases where there are alternative ways to realize the given good outcome. For example, suppose the altruist does not donate her pint, but does succeed in convincing the government to help the men in the desert. In that case, it is fine if the account gives us a reason to donate *or* to do something that realizes the good outcome all by itself.⁸ The point is: it’s *not* fine if the account gives us a reason to donate one’s pint *or* to convince the government *or* to pour the pint onto the ground *or* to do none of these things. The latter two courses of action help no one.

Second point: the notion of *superfluity*. We will assume that X is superfluous with respect to some outcome O when your act won’t make a

⁸ Such alternative ways may also face collective action problems, e.g. if 1000 people are needed to convince the government. For simplicity, in the cases to follow we will assume that there is only *one* course of action leading to the good outcome.

difference as to whether O will occur.⁹ Whether or not it is still up in the air what will happen, you know that either enough people are joining and O will be realized regardless of your efforts (e.g. enough people are donating their pint to the cart, and the men are helped regardless of your donation), or that there are not enough people joining and O won't be realized (even with your contribution there is not enough water to save the wounded men).

Third point: in the collective actions we focus on, *all* acts are superfluous and fall under the foregoing description. There is no individual of which we can say: if *that* person would keep her pint or steal the bean, some wounded man or tribesman would suffer more.¹⁰ No one makes a difference for the better or the worse. For the accounts that we will discuss, this is fine. For these accounts, it does not matter if X is superfluous, the idea is, because one may have a reason to do X nonetheless. Next, we will discuss two such accounts that utilize the notions of expression and integrity. We will argue that both are undermined by the superfluity problem, and this will illustrate and strengthen the challenge posed by it.¹¹

⁹ We are working here with the definition that Nefsky (2015: 264) uses to set up the superfluity problem. Ultimately, Nefsky (2017) herself drives a wedge between “non-superfluity” and “difference making.”

¹⁰ In triggering cases, there is a small chance that one does make a difference at the time of acting (e.g. that the election of a certain candidate turns on one's vote), and then, as Kagan (2011) argues, one may have an expected utility-based reason to act. In this paper, we will focus on non-triggering cases, i.e. where an act *could not* make such a difference, and so where expected utility calculations cannot be applied (cf. Nefsky 2011). Even so, the accounts that we will discuss *also* apply to triggering cases, and may provide additional reasons besides those of expected utility (as we will discuss in §5).

¹¹ The Undercutting Counterexample to versions of these accounts are also discussed in Nefsky (2018: 275-83). Our aim, here, will be to briefly summarize these worries, alongside those posed by the Alternatives Counterexample. Nefsky's claim, in that paper, is not so much that these accounts fail, but that they presuppose an “independent” solution. E.g. boycotts might still carry expressive significance when it is already established that people have an independent reason to perform them. Our aim will be to show that the participation account might well provide such reasons (and we will contrast this with Nefsky's own solution).

3. Undermined accounts: expression and integrity

First: the expression account. On this account, performing an action is not primarily about changing the world, but about making a statement, about expressing that one stands for something.

Expression. S has a reason to do X if by doing X, S expresses her disapproval of the given bad outcome (and her approval of the related good outcome).

For example, by donating her pint, the altruist expresses her disapproval of the suffering of the wounded men, and by refraining from stealing the bean the bandit expresses her disapproval of stealing from the tribesmen.

The precise details of the account depend on at least two things. First, there is a relevant distinction between signals one believes one is sending and the signals others believe one is sending. You may not *want* to send any signal, but if you steal beans or donate pints others might interpret this in a certain way.

Second, the account can be read in a consequentialist way or not. On a non-consequentialist reading, expression is valuable in itself, i.e. irrespective of any consequences this might have. On a consequentialist reading, by contrast, it is about the consequences of one's signals and whether one can motivate others to join the group (e.g. of altruists who donate their pint) and make a difference in this sense.¹²

Irrespective of such further details, we think the expression account has trouble with both counterexamples. Undercutting Counterexample. Donating a pint is superfluous with regard to the situation of the people in the desert. Suppose one knows this. Why would donating one's pint count as expressing disapproval of the suffering of the people in the desert (and approval of the relief of their suffering), if you know that donating is merely superfluous? Similarly, why should other people think that one cares about those people in such a case, and be motivated to start donating their pints?

¹² For two recent proponents, cf. Lane (2018) and Lawford-Smith (2015, 2018). Lawford-Smith suggests that by boycotting sweatshop products, one sends a signal that one disapproves of the sweatshops, and that one is willing to form a group that can make a difference for the better.

Alternatives Counterexample. The altruist could instead hold up a banner “DONATE YOUR PINT” in a public place all day. Holding up such a banner also counts as an expression of disapproval of the men’s suffering. Moreover, given that the altruist does so in public, other people will think she disapproves of this, and this may motivate them to hold up banners as well. But, if many people start holding up banners rather than donating their pints, then the men in the desert won’t be helped.

One might wonder whether this expression is sincere. That is, does one make a sincere statement that one disapproves of the men’s suffering if one does not donate one’s own pint? We think that the expression might well be sincere, namely if one sincerely believes that donating one’s pint is superfluous, and that one has no reason to do something superfluous.

The second account we will consider is the integrity account. On this account, it’s again not primarily about changing the world, but about being true to oneself, about acting in line with one’s own commitments.

Integrity. S has a reason to do X if by doing X, S avoids compromising her integrity.

For example, by donating her pint or by refraining from stealing the bean, the agents won’t compromise their integrity. Moreover, these agents *would* compromise their integrity if they were to keep their pint or steal the beans.

Undercutting Counterexample. The altruist has a commitment to help people in need. Donating her pint is superfluous, and the altruist knows it. Because it is superfluous, donating her pint won’t help others in need. Therefore, not donating her pint does not contradict her commitment to help people in need and so it does not compromise her integrity. Here, the integrity account fails to explain that the altruist has a reason to add her pint.

One might suggest that even though declining to donate one’s pint does not contradict one’s commitment to help people in need, it might still contradict *other* commitments one has. For example, one might contradict a commitment to participate in collective actions or in groups that might make a difference (regardless of whether one’s individual conduct will make any difference). Even so, it seems implausible that many people have such a

commitment, and in that case the integrity account won't yield reasons for action for such agents.¹³

And that's just one problem. Alternatives Counterexample. Again, the altruist could also, rather than donating her pint, hold up a banner in a public place. Starting such a campaign is in line with her commitment to help the wounded men (or at least her commitment to carry out acts that might help them). But, as before, if many people hold up banners rather than donate their pints, the wounded men won't be helped.

This account can again be read in a consequentialist or non-consequentialist way. On a non-consequentialist reading, integrity is valuable in itself, i.e. irrespective of any consequences it might have. On a consequentialist reading, compromising one's integrity is bad because it weakens one's disposition to act morally in the future.

Moreover, the same counterexamples may be generated in consequentialist terms. First, the altruist can do other things to avoid weakening her disposition, for example, by posting reminders in her living room. Second, if the altruist knows that donating a pint is superfluous, and she is sophisticated enough, then not donating would not weaken her disposition to act morally. Even so, in our view this agent is mistaken about what morality requires of her: the altruist does have a reason to add her pint, even though it is superfluous. And we are still looking for an account that can deliver such reasons.

4. Participation

Thus far we have seen that the superfluity problem is a serious problem that afflicts quite a few accounts, and this makes one wonder whether there is any adequate solution to it. In the following, we will present our solution in steps, and start with discussing current takes on participation that we think won't solve the problem.

¹³ In response, one might suggest that we should change our definition of what compromising one's integrity means. Substantive accounts of integrity identify specific commitments one *should* have, and one could suggest that one should have a commitment to participate in groups that can make a difference. Yet: what does it mean to be such a participant? This is just what we set out to account for in §§4-5.

Generally, on participation accounts what matters is not whether one as an individual can make a difference and change the world, but whether one is part of a group that can do so. We will start with the following account suggested by Parfit (labeled “C7”):

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. Similarly, even if some act benefits no one, it can be what someone ought to do, because it is one of a set of acts that together benefit other people. (1984: 70)

In terms of reasons for action, the account is this:

Participation. S has a reason to do X if by doing X S participates in a group that benefits others, and S has a reason to refrain from doing X if by doing X participates in a group that harms others.

For example, by adding her pint the altruist participates in a group that benefits the wounded men, and by stealing the bean the bandit participates in a group that harms the wounded men.¹⁴

For the account to work and form a solution to the problem set out at the start, it should accept that one can participate in a group that realizes a certain outcome *even when* one’s participation makes no difference to that outcome. Yet, this is precisely what leads to the Undercutting Counterexample. Stealing a bean is superfluous with regard to harming the tribesmen. The participation account grants this. But if stealing a bean is superfluous, then why should it count as “participating in a group that harms the tribesmen”? As Nefsky says: “It is unclear ... why I should count as part of the group that harms if what I do makes no difference with respect to that harm.” (2015: 249)

¹⁴ What if there are not enough donors to help the wounded men? Or what if there are not enough bean stealers to harm the tribesmen? We will address these issues in §5.

Whether this problem can be handled, Nefsky notes, depends on what precisely it takes to participate in a group. Parfit himself acknowledged this problem:

In my appeal to the effects of what groups together do, I did not fully explain whom we should count as members of these groups. This gap needs to be filled. There are many actual cases where, because the effects of each act are imperceptible, it is hard to tell who are the members of the groups who together harm or benefit other people. (1986: 848)¹⁵

But he goes on to suggest:

Often, however, we *can* tell. Thus it may be clear who are the people who together cause pollution, or congestion, or soil erosion, or the depletion of many kinds of resources. (ibid)

This won't work. For why would it be clear, for instance, who is in the group that causes pollution? That is, *according to what criterion* is it clear who's in the set, and who's out? Nefsky, in turn, considers and rejects Kutz's membership criterion, and then moves on to discuss other accounts not based on the notion of participation. In the following, we will expand her discussion and examine further possible membership criteria.

Kutz's membership criterion is as follows: S participates in a group that causes O iff S acts with an intention to do her part ("a participatory intention") in joint project P, and P causes O (Kutz 2000: ch. 3). For example,

¹⁵ According to Parfit (1986: 848), we cannot say: X is a member of the set that causes O iff X makes an imperceptible contribution to O. For, as he briefly notes, imperceptible harms and benefits involve Sorites-style worries: if one (additional) contribution is imperceptible, then 1+1 contributions are imperceptible, 2+1 contributions, and n+1 contributions will always remain imperceptible, and so will 999+1 contributions. But the latter is absurd, for we know that 1000 contributions will help the wounded men. Additionally, such a membership criterion does not capture what is distinctive about participation. If one admits imperceptible harms, one can altogether skip participation, and directly say that S has a reason to refrain from doing X if doing X causes imperceptible harm (cf. Barnett 2018).

if the altruist donates with the goal of doing her part in the joint project of helping the wounded men, and the men are being helped, then she belongs to the group of helpers, and if she does not act with that goal, she is not in that group.

Nefsky (2015: 251) notes that this account has difficulties with unstructured groups. For example, people who keep their pints do not usually act with the goal of *not* helping the wounded men (but rather with the goal of enjoying the pint themselves), and then the participation account does not seem to generate a reason for them to donate. Irrespective of whether such additional problems can be handled, we would like to show that Kutz's account also faces the two counterexamples.

Undercutting Counterexample. The altruist is sophisticated enough to know that adding her pint is superfluous. Hence, she reasons, why should adding my pint count as benefiting others? Given that she does not see this, she does not act with the goal of doing her part in the joint project of helping the wounded men. In such a case, she would not count as a participant, on Kutz's criterion, and the account fails to generate a reason for her to donate.

Alternatives Counterexample. Suppose you put a fresh coat of paint on the water cart. You hope this will attract the attention and have the effect of getting enough people to donate their pints. Here, you do act with the goal of doing your part in the joint project of benefiting the wounded men, and participate in this way. However, as we have said, we need a reason specifically to donate one's pint. Hence, Kutz's account won't help to block the counterexamples.

At one point, Parfit (1984: 71) suggests the following membership criterion (labeled "C8"): S participates in a group that causes O iff S is a member of the *smallest* set for which it is true that O wouldn't have occurred if its members had acted differently. Consider a variant of Harmless Bandits in which just two bandits steal the lunch of one tribesman. Each takes 500 beans from the plate. In that case, we want to say that these two bandits participate in the group that cause harm, and no one else. But now consider Fred Astaire who is just making a painting of the village. It still holds that the harm would not have occurred had the two bandits *plus Fred* acted differently. But Fred should not be part of the group that steals the lunch, and that is why the criterion speaks of the "smallest set". The smallest set for which the criterion holds, in this case, consists of the two bandits.

However, this criterion is still susceptible to the Undercutting Counterexample. In terms of Harmless Bandits it reads: the bandit is part of the group of agents who harm the tribesmen if she is a member of the smallest set for which it is true that the tribesmen would not be harmed if its members had acted differently. The problem is that, in collective actions like Harmless Bandits, there are no smallest sets for which this holds. After all, stealing a bean is superfluous, and such an act will never make a difference between a group being big enough for causing harm and a group being too small. But if there is no smallest set, then the bandit cannot be member of it, and so, on the given criterion, she is not part of the group of agents who harm the tribesmen.¹⁶

For the same reason, we cannot use the NESS-condition (widely employed after Hart & Honoré 1985: 112-3). In terms of this condition, S is a member of the set that causes O iff S's conduct is a necessary member of a set of acts that's sufficient for O. The idea is that S's conduct need not be necessary for O, but only for *some* set sufficient for O. For example, stealing a bean is need not be necessary for harming the tribesmen, but only necessary in *some* set sufficient to cause the harm.

To illustrate, consider again a variant of Harmless Bandits in which just two bandits are sufficient to steal the lunch of one tribesman and harm the latter. If there is just one bandit, who cannot eat all beans, there will still be plenty of beans left for the tribesman. As it happens, a group of three bandits descend on the village, and together they finish the man's lunch. In this case, we cannot say that any of the bandits was necessary for harming the man. After all, had any of them stayed home, the other two bandits would still have finished the lunch. Even so, the bandits are still necessary in a different way. For there are three sufficient sets consisting of two bandits—{b₁, b₂}, {b₁, b₃}, {b₂, b₃}—and the members of these sets are necessary for the sufficiency of the set: if any of them would stay home, the tribesman would not be harmed by the sets of which that bandit is a part. For

¹⁶ Parfit's subsequent principles "C10" and "C12" suffer from a similar problem. For if donating one's pint is superfluous, the wounded men are not benefited *less* if an altruist decides not to donate (and their clauses (4) fail) (cf. Gruzalski 1986: 779-82, Otsuka 1991: 142).

example, if b_1 stays home, the tribesman would not be harmed by the sets $\{b_1, b_2\}$ or $\{b_1, b_3\}$ (though he would still be harmed by $\{b_2, b_3\}$).

Unfortunately, the same reasoning fails to apply to the original version of Harmless Bandits, or any other collective action involving numerous agents. In such cases, we may identify sufficient sets. We could say: these 855 bandits are sufficient for harming the tribesmen, and these 767, and these 999. Thus there are many sets of acts sufficient for harming the tribesmen. But we cannot identify necessary elements of those sets (just as we could not identify the smallest sets before). For, stealing a bean is superfluous, and so if a set of size 767 is sufficient, so is a set of 766 bandits. Or if 855 bandits are sufficient, then so are 854. Generalizing, we can say that for any sufficient set of size n , $n-1$ would also have been sufficient. Hence, for any group sufficiently big to cause harm, no single act will ever be necessary for the sufficiency of that group being big enough for causing harm.¹⁷

Are *all* criteria doomed to fail for similar reasons? We think not. Basically, what we need is this. We need a criterion according to which any bean stealer is part of the group that causes harm to the tribesmen, but where Fred—who is just making a painting of the village—is not part of that group. That is, the group that harms the tribesmen in Harmless Bandits are all and only the bean stealers. Similarly, the group that benefits the wounded men in Drops of Water are all and only those who donate their pint. In the next section, we will make this criterion more precise, before showing how this blocks the superfluity problem.

5. Our account

Here's a first general statement of our membership criterion:

¹⁷ This problem afflicts the non-triggering cases we are considering here. In triggering cases, there *are* necessary members of sufficient sets (cf. Barry & Øverland 2016: 231-2). E.g. if there are 1000 voters needed to elect a certain candidate, then the members in a set of 1000 voters can still be considered necessary in the relevant sense.

(P1) S is a member of the set that causes O iff S does the act X which is such that: because enough others do that kind of act, they together cause O.

The bandit is part of the group that harms the tribesmen because she steals beans and this act is of a type that is such that, because enough others do acts of that type, they together harm the tribesman. The altruist is part of the group that benefits the wounded men because she donates her pint and this act is of a type that is such that, because enough others donate as well, they together benefit the wounded men. Does this proposal work? Parfit thinks not:

We cannot simply count, as members of the relevant groups, all those who act in certain kinds of ways. In some cases, the members of the relevant groups do not all act in the same way; in other cases, those who act in the same way are not all members of these groups. (1986: 484)

Parfit makes two claims: acting in the same way is neither sufficient, nor necessary for membership. Parfit does not offer examples, but think about driving cars and an (aggravated) greenhouse effect. Since one can also drive cars on green energy, driving cars is not sufficient for being a member of the group that causes the greenhouse effect. Nor is it necessary. After all, one can be a member of the group that causes the greenhouse effect in many ways: some drive cars, others run polluting factories, still others keep cows, and so on.

Let us take these two claims in turn. In response to the claim that acting in the same way is insufficient for membership, we think that much depends on the specificity of the description. If we specify enough, and distinguish green energy drivers from petrol drivers, then “those who act in the same way” are all members of the group that causes the greenhouse effect. After all, given enough specification, green energy driving is a different kind of action from petrol driving, and however many people drive cars on green energy, they together will not cause a greenhouse effect. Hence, if the level of description is fine-grained enough, acting in the specified

way—driving petrol-fueled cars—is sufficient for being a member of the relevant group.

In response to the claim that many different kinds of action should belong to the group that causes the greenhouse effect (driving cars, running factories, keeping cows, etc.), and that performing the same action is not required, we would like to point out that there is still something common to all these actions. Namely, they all contribute to the same thing: the amount of greenhouse gases in the atmosphere. In other collective actions, there is a similar “underlying dimension.” Just as there is an amount of greenhouse gases in the atmosphere, there is an amount of water in the cart that people could add to.¹⁸ Furthermore, our proposal is that it is the addition to that underlying dimension that makes an agent a member of the relevant group.¹⁹ Our proposal reformulated:

(P2) S is a member of the set that causes O iff S does the act X which is such that: (i) X adds to an underlying dimension D, and (ii) because enough others add to D, D causes O.

This account does not assume that, by doing X, S makes a difference to O. Rather, it proceeds in two steps. First: S can make a difference to D. Even if one can make no difference to the suffering of the tribesmen or the wounded men, one can still make a difference to the amount of stolen beans, or to the amount of water in the cart. Second, D can make a difference to O so long as enough people add to D. The amount of water in the cart can make a difference to the suffering of the wounded men so long as enough people donate their pint.

¹⁸ Some of them do not seem to be about *adding* to the underlying dimension, but about *subtracting* from it. E.g. there is an amount of beans on the plates or an amount of fish in the sea that people can subtract from. Yet this is just a matter of terminology: when the bandits subtract from the amount of beans, they “add” to the amount of *stolen* beans, and we will use this technical rendering of the term.

¹⁹ Kagan (2011: 117) and Nefsky (2011: 375-6) identify such an underlying dimension, though do not employ this concept to account for membership, as we are doing here.

(P2) is a good, basic statement of our overall account, yet we think that its timeframe is ambiguous.²⁰ At the time of acting, it is often unclear whether a certain outcome will be realized. For example, at the time of collecting water, it is unclear whether there will be enough donors so that the wounded men will be helped. It is unclear, that is, if there will be a set that benefits. But if this is so, it is unclear whether one will be a member of such a set, and then, on the participation account, it is unclear whether one has a reason to donate. But that's a bad result, for the account should be able to generate such reasons, even when it is still unclear whether there will be enough donors. After all, we want to solve the collective action problem, and provide reasons that are such that if enough people act on them, certain outcomes are realized rather than others (see §2). From this forward-looking perspective, our account is as follows:²¹

(P3) When it's still open whether O or ~O will occur: S has a reason to do X if (i) X adds to D, and (ii) if enough others would add to D (which might or might not happen), then D will cause O (which might or might not happen).

For example: at the time of acting, the altruist has a reason to donate if (i) the act of donating adds to the amount of water in the cart, and (ii) if enough others would donate their pint to the cart (which might or might not happen), then the amount of water will benefit the wounded men.

What's important, on (P3), is that one can have a reason to donate even if it's unclear, or still open, that the group will be big enough to help the wounded men. The phrase "it's still open whether O or ~O will occur" can be taken in different ways. It may mean that it is not yet determined, by past facts together with the laws of nature, that enough people will join the group, though this is not a reading that seems relevant in this context. We don't want

²⁰ Parfit's "C7" (cited in §4) is ambiguous in a similar way, though a natural reading of "it is one of a set of acts that together harm or benefit" is that certain groups have already been formed and outcomes already been settled.

²¹ Here and below, "O" refers to the good outcome (the wounded men are helped), "~O" to the related bad one (the wounded men aren't helped). Also, we will combine the membership criteria with the participation account to see how it generates reasons for action.

to commit ourselves to any specific interpretation, though think that an epistemic reading is natural. On such a reading, it might mean that people are still gathering and collecting water, and that there is not enough information about whether enough people will join (perhaps some have not even decided yet). It might also mean that the group of donors is *in fact* already big enough to help, but that one does not have enough information about this.

In a different paper, Nefsky (2017) has suggested a kindred forward-looking account. Her proposal is that S has a reason to do X if by doing X, S helps bringing about O. When can one help? One can do this, Nefsky writes, when “it is possible that [O] will fail to come about due, at least in part, to a lack of X-ing” (2017: 2753). (P3) differs from Nefsky’s account in at least two important ways. First, for reasons explained, we do not define membership of the group of helpers in terms of the kind of act that the agents perform (“X-ing”), but rather in terms of their contributions to an underlying dimension. Second, and more fundamentally, Nefsky wants to offer an account of “helping”, while we are interested in participation. As we will discuss next, even if it is clear that enough or that not enough agents will join, and hence one won’t have helping-based reasons to act in certain ways, one might still have participation-based reasons.

As we see it, what there is reason to do, is a relevant question in itself, regardless of whether it can solve the collective action problem (e.g. whether it can provide reasons for enough people to donate so that the men are being helped). Therefore, one may also want to evaluate what people have reason to do *once* it is already set which outcomes will come about. For example, if it is set that the tribesmen in the village will be harmed to a certain given extent, then it is still relevant to know if you have a reason to refrain from joining the bandits. Specifically, there are two kinds of situation that we need to consider: when it’s settled that the good outcome “O” will be realized, and when it’s settled that the bad one “~O” will realized.

First suppose it’s settled that the tribesmen will be harmed. There are already enough hungry bandits descending on the village, and stealing their lunch. If you would join them, you would not make any difference to the extent to which the tribesmen would be harmed. The pain of hunger won’t be any different in virtue of a single bean more or less. In this case, our account will say that one still has a reason to refrain from stealing the bean because: (i)

stealing a bean reduces the amount of beans on their plate, and (ii) given that enough others do this, the tribesmen will be harmed.²² Stated generally:

(P4) When it's settled that $\sim O$ will be realized: S has a reason to do X if (i) refraining from X adds to D, and (ii) given that enough others add to D (which actually happens), D causes $\sim O$ (which actually happens).

Next suppose it's settled that the tribesmen will *not* be harmed. There aren't enough bandits to steal their lunch. In this case, our account will still say that one still has a reason to refrain from stealing the bean because: (i) stealing a bean reduces the amount of beans on their plate, and (ii) if enough others would do this (which actually does not happen), then the tribesmen would be harmed (which actually does not happen). Stated generally:

(P5) When it's settled that O will be realized: S has a reason to do X if (i) refraining from X adds to D, and (ii) if enough others would add to D (which actually does not happen), then D would cause $\sim O$ (which actually does not happen).

Our three principles (P3), (P4), (P5) are more precise versions of (P2), and so all employ the same membership criterion: one is part of the group *when one adds to the given underlying dimension*. Yet, these principles differ concerning the status of the group. According to (P4), when it's settled that the tribesmen will be harmed, one is part of the *actual* group that harms when

²² Compare Drops of Water, and suppose it is settled that the wounded men will be helped. There are already enough altruists donating their pint. In this case, (P4) will say that one still has a participation-based reason to donate (even though there might be expected utility-based reasons *not* to donate) because: (i) donating adds to the amount of water in the cart, and (ii) given that enough others do this, the men will be helped. Note that this case is slightly distinct from Parfit's case (1984: 82) where the cart is *completely* full and one can only add one's pint by simultaneously taking water out of the cart. There is *enough* water, but it is not full. If it were full, as Parfit also notes, why add one's pint in the first place? (P4) may not generate a reason to do this, for if the cart is full, the altruist could not add to D, but only change the way D is filled, and then its clause (i) fails.

one steals a bean.²³ According to (P3) and (P5), in contrast, when it's still open whether the tribesmen will be harmed or when it is settled that the tribesmen won't be harmed, one has a reason to refrain from stealing not because by doing so one will participate in any actual group that harms (after all, there is no such group (yet)), but because by doing so one will still participate in a *possible* group that harms, that is, a group in a possible world where enough others steal beans as well.²⁴ In (P5) the group that harms is merely counterfactually possible (it won't exist in the actual world), whereas in (P3) this group is still possible in the actual future (as far as one knows).

Nevertheless, (P3), (P4), and (P5) are all equally important parts of the account, and we do not assume that the foregoing differences between the three principles matter morally. That is, we do not assume that it is worse to participate in actual groups that cause harm than to participate in merely possible groups, or that it is worse to participate in possible groups that might still become actual than to participate in groups that might not (or that it is worse to not participate in actual groups that benefit than it is to not participate in merely possible groups, etc.). For, in all these cases, one participates in these various groups so long as one adds to the underlying dimension (e.g. the amount of stolen beans). This contribution is always actual, and this is what constitutes, we take it, the same participation-based reason to refrain from stealing in all three scenarios.²⁵

²³ In non-triggering cases, where many sets are sufficient for $\sim O$, one might think that we cannot speak of "the" actual set that causes $\sim O$. Our account may resolve this worry if "the" actual set causing $\sim O$ consists of all and only those who add to D.

²⁴ Barry & Øverland (2016: 237) suggest that what matters is whether X *could* be a member of the actual set causing harm. They do not offer specific membership criteria, and might well take our "underlying dimension" criterion on board. Even so, our account is different. Their account implies that if it is clear that X won't be in the actual set, one has no reason not to do X. As we say here, we think there are still participation-based reasons e.g. not to steal beans *even if* it is clear that the group of bean stealers is not big enough to cause harm (and there won't be an "actual set causing harm").

²⁵ That is: the same pro tanto reason. Thus, when it is still open whether O or $\sim O$ will occur, there may be *additional* reasons of helpfulness (for example) that might obtain that do not obtain when it is already settled that that O or $\sim O$ will be

Our account is quite liberal, and will generate reasons for many kinds of actions. To see this, let's contrast it to the expected utility account (favoured, most prominently, by Kagan 2011). According to the latter, what matters is whether the expected utility of acting is higher than that of acting otherwise, i.e. whether the benefits are expected to be higher than the costs compared to alternative acts. On our account, in contrast, what matters is whether you add to some underlying dimension D (which is such that, if enough people add to it, some difference will be made). Thus taken, our account is degreeless. If you add to D, you participate all the same. It doesn't matter, for example, (i) how likely it is that enough people add to D, or (ii) how costly it is for you to add to D.²⁶

Take factor (i), and suppose it is already settled that the tribesmen will be harmed. Numerous bandits are stealing their lunch. Now you could join them and take some beans for yourself, or not. In such a case, there is an expected utility-based reason to join them. After all, doing makes no difference for the worse (the probability of not harming the tribesmen remains zero), while it does give you a nice lunch. Even so, our account will say that one still has a participation-based reason against stealing in such a case. It is important to see the implications of this. Consider a comparable case by Sinnott-Armstrong:

Suppose that the president sends everyone (or at least most tax-payers) a check for \$600. If all recipients cash their checks, the government deficit will grow, government programs will have to be slashed, and severe economic and social problems will result. You know that enough other people will cash their checks to make these

realized. This would make the *overall* weight of moral considerations in the former case heavier than in the latter (but not the participation-based reason itself).

²⁶ Technically, it is still possible to make the account gradual by letting the strength of the participation-based reason depend on *other* factors such as: how much you add to D (one pint or six, say), or how good or bad the results are if enough people add to D (e.g. whether one saves 1000 persons or 3). See Lepora & Goodin (2013) for an elaborate account of such degrees (though these authors define membership in terms of "potential difference-making" (2013: 64-5), a concept that does not apply to non-triggering cases, where members cannot make a difference, not even potentially).

results to a great degree inevitable. ... In these circumstances, is it morally wrong for you to cash your check? I doubt it. (2005: 298)

In this case, one might think, like Sinnott-Armstrong, that it's permissible to cash the check. We agree that there are expected utility-based reasons to do so. Cashing will give you \$600, while you make no difference for the worse. Yet, this does not mean that there are no considerations that speak *against* your conduct, particularly participation-based reasons. For, by cashing the check, you will participate in the group causing the severe economic and social problems.

Next consider factor (ii), and suppose it is very costly to add your pint. Your friend is sick, for instance, and really needs your pint to get better. If you give it to your friend rather than to the water-cart, you fail to participate in a set that benefits the wounded men. On our account, there is still a participation-based reason against what you are doing, though this does not rule out *other* reasons speaking in favour of your choice. In this case, there might well be expected utility-based reasons to give the pint to your friend, namely if the expected benefits of this act are higher than that of donating it to the water-cart.

Hence, in many cases our participation account generates a certain reason for action, where the expected utility account does not. Even when the chance is really low that there are enough other boycotters to make a difference for the better (and the expected utility of boycotting is negligible), there are still participation-based reasons to refrain from purchasing sweatshop or animal products. In this way the account calls attention to considerations that are not acknowledged by the expected utility account (nor by Nefsky's account of helpfulness).²⁷ These may not be decisive reasons, but they are still *pro tanto* reasons, i.e. considerations that speak for or against certain courses of action.²⁸

²⁷ Yet, expected utility theorists may still add participation costs to the calculation (cf. Kutz 2000: 182-4). That is, it may not only be costly to donate your pint (it takes you e.g. time and energy), but also to keep it: you may feel bad if the wounded men die and you are not part of the group of altruists that tried to help them.

²⁸ Suppose 99% of people have voted for the good candidate, and voter turnout has been 45% with only a few minutes left in the polling. In the past, no more than 60% of people have ever voted. It is still open who will win, since every remaining

Moreover, that the participation account is liberal in this way is just what we should expect. At the start, we spoke about people who boycott sweatshop or animal products because they don't want to be complicit in harm. These people won't make their boycott dependent on what other people do. It does not matter if it is still up in the air if the boycott will be successful, or if it is already settled that it won't. What matters, instead, is what group they belong to. Even if it is already settled that there won't be enough other boycotters to make a difference for the better, these agents still don't want to belong to the group of agents who together sustain exploitation and animal suffering.²⁹

In the foregoing, we have defended a specific account of what it takes to be a member of such groups. As discussed, we do not think that membership is determined by, for example, the intentions of the agents, or the concerns their actions express. Instead, we have suggested a certain contribution principle according to which consumers are complicit when they contribute to the amount of sold sweatshop or animal products (which is such that, if this amount is big enough, more such products are being ordered, and more exploitation or animal suffering is being triggered).

6. Problem solved

In this paper, we have focused on a family of accounts according to which one can have a reason to take certain courses of action even if the latter makes no difference. Many accounts in this family are undermined by Nefsky's superfluity problem, or, more precisely, by what we have called

voter could vote for the bad candidate. On (P3), one still has a participation-based reason to vote for the good candidate. This may seem counterintuitive, though we think that this worry can be countered by acknowledging that we are concerned with pro tanto reasons. In the scenario described, there may well be an overriding expected utility-based reason *not* to vote. For it is certain there will be costs to voting (spending time and energy going to the ballot box) whereas the chance that your vote will make a difference seems too small in this scenario.

²⁹ These are the kind of agents that will refuse to buy chickens or other animal products even when they know *for sure* that they would otherwise go to waste.

“Undercutting Counterexample” and “Alternatives Counterexample.” Finally, we are able to show that our account blocks both counterexamples.

Undercutting Counterexample. Even if your act of stealing is superfluous with respect to harming the tribesmen, it may still add to the underlying dimension (the amount of stolen beans on the plates), and this makes you member of the (actual or possible) set that harms the tribesmen. Given that the participation account says we shouldn’t be in such sets, it yields a reason to refrain from stealing.

Alternatives Counterexample. Suppose you organize a huge awareness campaign. If your campaign is among a set of acts that ends the suffering, then the participation account says you should be in such sets. In any case, it’s also the case that if you fail to donate your pint, then you fail to add to the underlying dimension (the amount of water in the cart), which means that you fail to be a member of the (actual or possible) set that helps the wounded men. Given that the participation account says we should be in such sets, it yields a reason to donate. It may *also* generate a reason to campaign, but these reasons don’t conflict.³⁰

Thus, we think that Nefsky was right to criticize many accounts on the basis of the superfluity problem, though wrong think that no variant of the participation account would work. She writes: “Why can’t each individual say: my act makes no difference, and so it’s not part of the set of acts that together makes a difference for the worse?” (2015: 249) Individuals cannot say this, for even when their conduct is superfluous with respect to some outcome, they can still be said to add to the given underlying dimension.

As always, there are further things to say, and important issues are the following. First: why is participation morally significant and something to be avoided? Second: how weighty are participation-based reasons (as contrasted to, for example, expected utility-based reasons)? Now that we

³⁰ Whether it generates *too many* reasons is a familiar issue about reasons to benefit others, and poses no special problem for this account.

have shown that the participation account can answer to the two counterexamples, these follow-up issues have to be debated elsewhere.^{31, 32}

References

- Barnett, Z. 2018. No Free Lunch: The Significance of Tiny Contributions. *Analysis* 78: 3-13.
- Barry, C. & G. Øverland 2016. *Responding to Global Poverty. Harm, Responsibility, and Agency*. CUP.
- Budolfson, M. B. 2015. Is It Wrong to Eat Meat from Factory Farms? If So, Why? In B. Bramble & B. Fischer eds., *The Moral Complexities of Eating Meat*, ch. 5. OUP.
- Cullity, G. 2000. Pooled Beneficence. In M. J. Almeida ed., *Imperceptible Harms and Benefits*, pp. 1-23. Kluwer.
- DeGrazia, D. 2009. Moral Vegetarianism from a Very Broad Basis. *Journal of Moral Philosophy* 6: 143-65.
- Driver, J. 2015. Individual Consumption and Moral Complicity. In B. Bramble & B. Fischer eds., *The Moral Complexities of Eating Meat*, ch. 4. OUP.
- Glover, J. 1975. It Makes No Difference Whether or Not I Do It. *Proceedings of the Aristotelian Society* 49: 171-90.
- Gruzalski, B. 1986. Parfit's Impact on Utilitarianism. *Ethics* 96: 760-83.
- Harman, E. 2015. Eating Meat as a Morally Permissible Moral Mistake. In A. Chignell et al. eds., *Philosophy Comes to Dinner*, ch. 12. Routledge.
- Hart, H. L. A. & T. Honoré 1985. *Causation in the Law* (2nd ed.). Clarendon.
- Kagan, S. 2011. Do I Make a Difference? *Philosophy & Public Affairs* 39: 105-41.
- Kutz, C. 2000. *Complicity. Ethics and Law for a Collective Age*. CUP.

³¹ See Driver (2015) for two potential justifications. First, adding to D (in our terms) may be bad because it is due to certain objectionable attitudes. Second, adding to D may be bad in a rule consequentialist way: policies of not adding to D may well be optimific.

³² We thank: many colleagues and students at Vrije Universiteit Amsterdam for discussion, Holly Lawford-Smith, Anne Polkamp, Phil Robichaud, and two reviewers of the journal for written feedback, and the Talent Fund of the Clue+ Research Institute for financial support.

- Lane, M. 2018. Uncertainty, Action and Politics: The Problem of Negligibility. In K. Forrester & S. Smith eds., *Nature, Action and the Future*, ch. 8. CUP.
- Lawford-Smith, H. 2015. Unethical Consumption and Obligations to Signal. *Ethics & International Affairs* 29: 315-30.
- 2018. Does Purchasing Make Consumers Complicit in Global Labour Injustice? *Res Publica* 24: 319-38.
- Lepora, C. & R. E. Goodin 2013. *On Complicity and Compromise*. OUP.
- Martin, A. M. 2015. Factory Farming and Consumer Complicity. In A. Chignell et al. eds., *Philosophy Comes to Dinner*, ch. 11. Routledge.
- McPherson, T. 2015. Why I am a Vegan (and You Should Be One Too). In A. Chignell et al. eds., *Philosophy Comes to Dinner*, ch. 4. Routledge.
- Nefsky, J. 2011. Consequentialism and the Problem of Collective Harm: A Reply to Kagan. *Philosophy & Public Affairs* 39: 364-95.
- 2015. Fairness, Participation, and the Real Problem of Collective Harm. *Oxford Studies in Normative Ethics* 5: 245-71.
- 2017. How You Can Help, Without Making a Difference. *Philosophical Studies* 174: 2743-67.
- 2018. Consumer Choice and Collective Impact. In A. Barnhill et al. eds., *The Oxford Handbook of Food Ethics*, ch. 12. OUP.
- Otsuka, M. 1991. The Paradox of Group Beneficence. *Philosophy & Public Affairs* 20: 132-49.
- Parfit, D. 1984. *Reasons and Persons*. OUP.
- 1986. Comments. *Ethics* 96: 832-72.
- Singer, P. 1975. *Animal Liberation*. HarperCollins.
- Sinnott-Armstrong, W. 2005. It's Not My Fault: Global Warming and Individual Moral Obligations. *Perspectives on Climate Change* 5: 221-53.