

The Psychological Basis of Collective Action

James Fanciullo

Forthcoming in *Philosophical Studies*

Abstract: Sometimes, a group of people can produce a morally bad outcome despite each person's individual act making no difference to whether the outcome is produced. Since each person's act makes no difference, it seems the effects of the act cannot provide a reason not to perform it. This is problematic, because if each person acts in accordance with their reasons, each will presumably perform the act—and thus, the bad outcome will be brought about. I suggest that the key to solving this problem is to make it true of each person that their act would in fact make a difference to the relevant outcome. Fortunately, I contend, this can be accomplished by each person simply forming a particular type of attitude. I argue that each person has an obligation to form the relevant attitude in collective action cases, on pain of being immoral or irrational.

Let's begin with a story. Consider

Drops of Water: 10,000 wounded people lie out in the desert, suffering from intense thirst. We are an equally large number of altruists, each of whom has a pint of water. We could pour these pints into a water cart. This would be driven into the desert, and our water would be shared equally between all these many wounded people. By adding our own pint, each of us would enable each wounded person to drink slightly more water—perhaps only an extra drop. The effect of each of these extra drops on each person would be imperceptible.¹

¹ This is adapted from Parfit 1984 (p. 76).

If each of us adds our own pint, each of the wounded people will receive a full pint of water, resulting in the relief of each wounded person's suffering. It might seem obvious, in that case, that each of us ought to contribute our water to the cart. The problem, however, is that each of our individual contributions will make no difference at all to the suffering of these wounded people. The addition of our own pint would only enable each wounded person to drink an extra drop of water. Far from relieving anyone's suffering, this difference is not even great enough to be noticed. That being so, it seems that whether the suffering people receive enough water depends entirely on the contributions of others. Since adding our own water just isn't the kind of act that could have any influence on whether the suffering of these people is relieved, there doesn't seem to be any reason for us to add our water to the cart. So, while we might hope that a sufficient number of the other altruists add their own water, we have no reason to think that anyone with a pint ought to contribute.

This conclusion is likely to leave many of us dissatisfied. After all, if each of us were to employ the above line of reasoning, and we each acted in accordance with our reasons, no one would contribute their water to the cart. And if no one contributed their water to the cart, the suffering of those in the desert would not be relieved. While it seems clear that morality should have something more to say in such *collective impact cases*—or cases where acts (or omissions) of a certain type collectively bring about a bad (or good) outcome but individually make no difference with respect to that outcome—it is unclear what exactly this something more might be.² Since our act of adding water wouldn't make any difference to the suffering of those in the desert, it seems each of us acts permissibly by keeping hold of our water.³ Similarly, since our own carbon emissions won't

² I borrow this terminology from Julia Nefsky (2017, p. 2744).

³ Of course, some people deny that each contribution would make no difference. See, for example, Barnett 2018, Kagan 2011, and Norcross 2004. These people in particular argue that each individual contribution *might* make a difference. This approach faces serious difficulties, however, as Julia Nefsky (2011) and Mark Budolfson (2019) have shown.

make any difference to the suffering of future generations, and our own votes in national elections won't make any difference to who is elected, it seems each of us acts permissibly by both emitting at unsustainable rates and refraining from voting.⁴ This may of course lead us to shift our focus, in *Drops of Water* in particular, from the act of adding water itself to the psychological features of the agents performing it, but there is also no obvious error in reasoning or flaw of character to which we can attribute our collectively failing to reach the morally good outcome. Thus it seems we are left either to give up the plausible thought that collectively satisfying the demands of morality will not lead us to morally bad outcomes, or else to find some previously unrecognized psychological feature of those who do not contribute in virtue of which they might be blameworthy.⁵

In this paper, my aim is to introduce and explore a psychological feature of this sort. As we will see, closer inspection of cases like *Drops of Water* reveals the particular feature of the non-contributor's psychology in virtue of which she should be blamed. Before getting to these points, though, it will be helpful to consider a number of views that try, and fail, to generate a plausible reason for us to contribute in cases like *Drops of Water*.

1. A few apparent solutions

The thought that morality must have something more to say about cases like *Drops of Water* is, of course, not new. A number of attempts have been made to show that there is some reason, issuing from morality, to contribute one's water to the cart. Still, while these suggestions may initially seem plausible, each is problematic in its own right.

⁴ While these real-world cases may not be perfectly analogous to *Drops of Water*, they are nevertheless collective impact cases, and so a solution to our problem in *Drops of Water* would also seem to apply to them.

⁵ Taking the first option would also force us to reject the stronger claim that the demands of morality will not lead us to morally suboptimal outcomes, or (roughly) what has elsewhere been called "the principle of moral harmony." Proponents of this principle include Parfit (1984, p. 54), Pinkert (2015), Portmore (2018), and Regan (1980).

The first of these views claims that, if we do not add our water to the cart, we are blameworthy in virtue of our having failed to care about the suffering people to a sufficient degree. On this suggestion, “to care about X is to be disposed to react and act in ways conducive to its going well with X in some regard” (Björnsson and Brülde 2017, p. 17).⁶ So, according to this view, one is required to care about those in the desert at least to the extent that one is disposed to add one’s water to the cart. Failing to add one’s water reveals that one lacks such a disposition, and it is in virtue of lacking this disposition that one is blameworthy. Accordingly, the reason that one is blameworthy for failing to add one’s water, on this view, is that one did not have the disposition to act in a way conducive to relieving the suffering of the people in the desert. Possession of this disposition is, we might say, a minimal requirement for caring, on the current view.

There are at least two problems with this proposal. The more fundamental problem is that the view simply stipulates that we are required to have a certain disposition concerning the wounded people, and offers no reason for us to think that this requirement is plausible. The view is then no different from one claiming that the reason we have to add our water is that we are morally required to do so. Such a view would simply ignore the problem with which we are concerned, telling us that we are required to act in some way, but saying nothing about why we are so required. The problem in *Drops of Water* in particular is that given our act of adding water would do nothing to help relieve the suffering of any wounded person, it is unclear how we could have any reason at all—let alone be morally required—to contribute our water to the cart. To claim that our reason to add our water to the cart is that we are morally required to do so is thus to ignore the problem altogether. Which brings us to a second, related issue for this view. The proposed view tells us that we ought to be disposed to add our water to the cart, because adding our water is acting in a way conducive to

⁶ It is worth mentioning that these authors do not present their account of caring as any kind of solution to cases of the sort we’re currently considering. I use their account only because I see the current proposal as worth considering, given its potential initial plausibility, and see their account as a good one for filling the proposal out.

relieving the suffering of the people in the desert. The very point of contention, though, is whether we have any reason to add our water to the cart, given our act would make no difference to the suffering of those in the desert. And notice that this proposal is helping itself to the claim that to add one's water is to act in a way conducive to relieving their suffering. It seems, then, that the proposed view assumes that one's act of adding water does make a difference, and concludes on the basis of this assumption that one is required to be disposed to add one's water. This assumption is false, however, and the corresponding conclusion—which is little more than an assumption itself, as we have just seen—is unsupported. Consequently, it seems the current proposal should be rejected.

An alternative view tells us that, while no one of us has an individual obligation to contribute our water, we together have a collective obligation to add enough water to the cart.⁷ Since we as a collective can determine whether the suffering people will get enough water—an amount that, given any individual contribution would not be noticed, must be less than all 10,000 pints—we are collectively obligated to add enough water to ensure the relief of their suffering. Thus, given this obligation, each of us is obligated to contribute our own water; and if we fail to contribute, we are guilty of being a “moral free-rider”—or someone who excuses themselves from doing the work needed to ensure that the good outcome is brought about.⁸ Of course, since the number of pints needed to relieve the suffering of those in the desert is less than 10,000, not everyone must contribute in order to ensure that this good outcome is brought about. But, even so, the current view tells us that if we do not contribute, we act unfairly. So, we should each add our water to the cart, on this view, because we are under a collective obligation to get the suffering people a sufficient amount of water, and to add our own water is to do our part in satisfying this collective obligation.

⁷ See Cullity 2000. For objections to this view, see Nefsky 2015.

⁸ See Nefsky 2017 (p. 2748).

This view faces at least three worries, two of which should seem familiar.⁹ The first problem is, again, that the view seems to put the cart before the horse: it assumes that we have a collective obligation, and then concludes that we are blameworthy for failing to act in accordance with this obligation. But why think that we have a collective obligation at all? If I myself have no obligation to add my water, and the same is true of everyone else who has a pint, it is unclear how this collective obligation is supposed to arise.¹⁰ Thus, much like the first view we considered, this view assumes that we are required to add our water to the cart, but offers no justification for this assumption. The second problem for this view is that, considering I myself am under no obligation to add my water, it is unclear how it could be that by failing to add my water I become a no-good “free-rider.” To be sure, I’m excusing myself from the work of contributing, but I also could not have made any difference to the amount that the others contribute. There was nothing I could meaningfully add to their contribution, so why should we think that I’m free-riding, rather than merely unable to make any difference to the outcome? Hence even if there is a collective obligation, it seems permissible for each of us to refrain from contributing. The third and final problem concerns appeals to fairness. If the claim here is that, when everyone contributes except me, the unfairness consists in my getting to keep my water while everyone else gives up theirs, then it seems I could equally satisfy the requirements of fairness either by adding my water to the cart or by dumping my water on the ground. This means that, given a choice between keeping hold of my water and dumping it on the ground, the current view says to dump it on the ground. And it says this regardless of the fact that my water wouldn’t make any difference with respect to satisfying our collective obligation. This, I think, is implausible. Consequently, even if we are under some collective obligation—which this view simply assumes to be the case—it does not seem plausible that this would entail any

⁹ Nefsky (2015) also discusses this, as well as the following, failed solution to the problem of collective impact, and the following few paragraphs have greatly benefitted from her discussion.

¹⁰ See Tänssjö 1989.

corresponding individual obligations, and thus it would still be permissible for each of us to keep hold of our water. Accordingly, it seems this view, too, should be rejected.

One natural response to the preceding discussion is that we are too narrowly focused on the *consequences* of our act of adding water. If we move away from this consequentialist approach to the problem, however, we might find a solution. Consider, for example, the perfectly virtuous person, or perfectly dutiful Kantian. Surely each of these people would find a flaw in the consequence-based reasoning employed above. It might be, for example, that the perfectly virtuous person would add her water out of sympathy for the people who are suffering.¹¹ Alternatively, perhaps the perfectly dutiful Kantian would add her water to avoid a contradiction in her will. Either way, on this suggestion, we seem to have a reason to add our water to the cart, where this reason has nothing to do with the consequences of our individual action.

Notice, however, that to entirely ignore the consequences in a case like Drops of Water would be, not shrewd, but irresponsible. To see this, consider first the perfectly virtuous person. Even if adding her water to the cart would display sympathy for the people who were suffering, her act would be no more than that—a display.¹² After all, adding her water to the cart wouldn't make a difference to their suffering. So why think that she would in fact add her water, rather than dismiss the act as wasteful? So long as adding her water to the cart makes no difference to anyone's suffering, she has no reason to think contributing is conducive to virtue and failing to contribute is

¹¹ This type of approach has been adopted by a number of philosophers with respect to individual action in the face of global environmental change. See, for example, Hill 1983 and Jamieson 2007.

¹² It might be argued, of course, that the act may still be good to perform, even if it would be no more than a symbolic display. See, for example, Hill 1979. On this type of view, by adding our water to the cart, we disassociate ourselves from a policy that would lead to a bad outcome, and thereby honor a type of cause that is more morally deserving (Hill 1979, p. 99). The problem here, however, is that it simply isn't clear that by adding our water we are associating ourselves with any cause at all. Since adding our own water would not make any difference to the outcome, it's unclear how this act would constitute our associating ourselves with a good outcome or policy. Again, we need a view that explains why our act would be a good one to perform; and so long as this act would make no difference, it will not help to simply paint our act as one that would be good to perform, as this view (among others) would seem to do.

conducive to vice. Hence, it seems, the appeal to the perfectly virtuous person fails to reveal any reason for us to contribute our pint in Drops of Water.

Consider next the perfectly dutiful Kantian. Surely this person would take into account the fact that adding their water would have no effect on the suffering of those in the desert. Thus it doesn't seem that by failing to contribute this person would use anyone else as a mere means to their end. Nor does it seem that the maxim of their action would entail a contradiction in conception. If the maxim of this person's action was, for example, "don't contribute when doing so would make no difference to the relevant outcome," there would be no contradiction if everyone were to adopt the same maxim. Of course, it might be suggested that this person nevertheless has an *imperfect* duty to help the suffering people, in which case she has at least some reason to add her water to the cart. But notice that this response only works so long as adding one's water to the cart would in fact help the suffering people—or, so long as this act would make a difference to their suffering. Since adding one's water would make no difference to anyone's suffering, this response fails. Accordingly, the appeal to the perfectly dutiful Kantian also fails to reveal any reason for us to contribute our pint in Drops of Water.

In sum, finding a plausible view on which non-contributors in cases like Drops of Water may be blameworthy is not easy. None of the approaches we've considered seem promising.¹³ A common problem with these approaches is that, while they establish a reason for us to contribute, they do so only by assuming that we are obligated to contribute. The reason that this problem is particularly common is, I think, that given our distaste with the conclusion that we have no reason to act in cases like Drops of Water, we are intuitively compelled to the thought that there must be

¹³ In a recent paper, Nefsky (2017) has argued that each person does have a reason, issuing from morality, to add their water to the cart. This reason stems from the fact that, while each act cannot make a difference, each act could *help* relieve the suffering of those in the desert—that is, each act could make a non-superfluous causal contribution to bringing the morally good outcome about. I think that this approach is inadequate as well, however, for reasons offered in Fanciullo 2019.

some obligation for us to act. This leads some to assume that there is such an obligation, and to conclude on this basis that we do indeed have a reason to act. This approach, we have seen, is not promising. But this does hint at a more promising way forward. If what leads us to assume that we are under some obligation to contribute is our distaste with the conclusion that we have no reason to contribute, examining this latter intuition may prove useful. In particular, a closer look at this intuition might help us find a plausible reason to contribute in cases like Drops of Water.

2. Isolating our intuition

Most of us, I suspect, have the intuition that there must be some reason for us to contribute in Drops of Water. My aim in this section is to determine which features of Drops of Water are relevant to this intuition. The hope is that our investigation here will help us locate a plausible reason to contribute in cases of this sort.

There is something counterintuitive about the conclusion that the suffering people might receive no water, yet each of us has acted permissibly by keeping hold of our pint. In order to determine what, exactly, this intuition is tracking, we can shift particular features of the case and see whether our intuition changes. So consider

One Mechanism. Your situation is just as it is in Drops of Water, except now there are no other altruists. Instead, there is a mechanism that will either release 9,999 pints of water into the cart, or it will not. You are just as confident that the mechanism will add the water as you are that, in Drops of Water, enough other altruists will add their water to ensure the relief of the people in the desert's suffering.¹⁴

¹⁴ Each of the variants of Drops of Water considered in this paper are introduced in Fanciullo 2019.

What is our intuition in this case? My guess is that almost everyone will agree: you have no reason to add your water in One Mechanism. After all, whether the good outcome is brought about seems to depend exclusively on whether the mechanism releases the water. Your pint would never be noticed if it were added to 9,999 others, and your pint alone would do nothing to relieve the suffering of the many people in the desert. Either way, then, your act seems pointless.

Notice, though, that your act in One Mechanism seems pointless in a way that separates the case from Drops of Water. In the latter case, while we may be *brought* to conclude that we have no reason to add our water to the cart, this is not the intuition with which we begin. In One Mechanism, on the other hand, the thought that we have no reason to add our water to the cart seems intuitively clear. This is true despite the fact that the odds of the good outcome being brought about are in each case the same. The first thing for us to notice, then, is that the difference between our intuitions here cannot be explained by a difference in the likelihood of the many wounded people getting enough water.

Then what is the difference between these two cases in virtue of which our intuitions about them change? Well, in moving from Drops of Water to One Mechanism, only two features of the case change. And so, it will presumably be at least one of these two changes that account for our shift in intuition. The first, perhaps more conspicuous change concerns the number of potential contributors: in Drops of Water there are 10,000, and in One Mechanism there are two. It might be, then, that the reason your potential act of adding water seems truly pointless in One Mechanism, but not Drops of Water, is that in One Mechanism there is just one other potential contributor, whose contribution would unilaterally determine whether the group's suffering was relieved. So perhaps the fact that there is just one other potential contributor, whose potential contribution is much weightier than our own, explains our intuition about One Mechanism.

I do not think that this is right. To see why, consider another case:

9,999 Mechanisms: Your situation is just as it is in *One Mechanism*, except now there is more than just one mechanism. Instead, there are 9,999 mechanisms, each of which will either release a pint of water into the cart, or it will not. You are just as confident that enough mechanisms will add their water to the cart to ensure the relief of the people in the desert's suffering as you are that, in *Drops of Water*, enough other altruists will add their water.

Intuitions here may diverge, but let me report mine: I think that it is only slightly less clear that you have no reason to contribute in 9,999 Mechanisms than it is clear that you have no such reason in *One Mechanism*. The fate of the suffering people depends on events that are wholly outside of your control: the odds will either play out in a way that ensures the relief of the people's suffering, or they will not. You might hope that they ultimately play out in the first way, but there is nothing that you can do to make this outcome more likely. Just as in *One Mechanism*, then, it seems your act of adding water here would be truly pointless. In this case, though, there are just as many mechanisms as there are other altruists in *Drops of Water*. Thus, so long as our intuition about whether you have any reason to contribute in 9,999 Mechanisms aligns with our corresponding intuition about *One Mechanism*, it seems the number of potential contributors is not the variable that we are looking for. This is because, if contributing seems pointless—or, at least, to have less of a point—even when the number of contributors is equal to the number of altruists in *Drops of Water*, it would have to be some other feature of the cases that differentiates them in the relevant way. Accordingly, the difference between *One Mechanism* and *Drops of Water* in virtue of which our intuitions about the cases change does not appear to be that there are fewer potential contributors in the former.

It is perhaps worth mentioning that, as soon as I see that *One Mechanism* and 9,999 Mechanisms are in all relevant respects the same, I have little problem endorsing the conclusion that

we have no reason to contribute in either case. But the same is not true of Drops of Water. I am left dissatisfied with the conclusion that we have no reason to contribute in Drops of Water in a way that I am not dissatisfied with the same conclusion in these other cases. This, too, suggests that the difference in the number of potential contributors is not responsible for the change in our intuitions about Drops of Water and One Mechanism. After all, if it were so responsible, we would presumably have identical intuitions about Drops of Water and 9,999 Mechanisms. But, I have suggested, we do not. Thus, the difference in number of contributors does not seem to be the difference that we are looking for.

This leaves us with just one difference between One Mechanism and Drops of Water that might explain the difference between our intuitions about these cases. This difference concerns the identity of the other potential contributors: in Drops of Water they are agents, and in the mechanism cases they are mechanisms. It might be, then, that our distaste with the conclusion that you have no reason to contribute in Drops of Water derives from the fact that other people are in the same situation as you, and can thus reason their way to a similar conclusion. I do not mean to claim, though, that the explanation of our distaste here is that, if we apply a “what if everyone did that?”-type universalizability test, we would reach a bad outcome. As we have seen, at least one version of this approach fails to locate a reason for us to contribute; and arguments have been provided elsewhere for thinking that other versions of it similarly fail.¹⁵ Rather, the explanation of our distaste is that, given these people are in the same situation as you, it might also be possible for each of you to reason your way to a different conclusion—one that results in the suffering people getting enough water. So perhaps the reason your act of adding water does not seem truly pointless in Drops of Water is that, under certain circumstances—circumstances that involve reasoning in

¹⁵ See Nefsky 2015 and Sinnott-Armstrong 2005.

concert with others—it is possible that your act isn't pointless. That is to say, perhaps under these circumstances, it's possible that your act makes a difference.

This suggestion is, of course, little more than speculation. In fact, it seems at this point more like a guess. But even so, the approach seems to me promising. After all, if we can find some way to make your act such that it would make a difference, a reason to perform the act will naturally follow. Given that our intuition about Drops of Water seems to depend importantly on the fact that the other contributors are agents, it seems plausible that this fact may lead us to our desired conclusion. So, perhaps there is a way to make your act such that it would make a difference, where this method is available only in virtue of the fact that agents are the other contributors. Let's see how far this plausible thought can be taken.

3. How we each can make a difference

As we've seen, the difference between 9,999 Mechanisms and Drops of Water in virtue of which our intuitions about the cases change seems to be that only in Drops of Water are the other potential contributors agents. My hypothesis, or at least the one that I would like to try out, is that our intuition about non-contributors in Drops of Water being worthy of blame derives from their failure to act in concert with the other agents in order to reach the good outcome. So while it is true of each altruist that, as described, they have no reason to contribute their own pint, it seems to us that something more could be done. Setting out precisely what I take this something more to be is the task of the current section.

First things first. Notice that it cannot be any altruist's act of keeping hold of their water *itself* that we take issue with here. We know that the alternative act, adding their water, would make no difference to bringing the good outcome about, and it therefore seems permissible for each altruist to keep hold of their water. In fact, given these people have no reason whatever to contribute,

rationality seems to *demand* that they keep hold of their water. So it cannot be the act of withholding water itself that we find problematic—this is the best option they have. The thing we take issue with, then, must relate to the altruists' *attitudes*.

The next thing to notice is that if we stipulate that the altruists are completely shut off from each other, and the good outcome is thus not produced, the thought that morality should have something more to say—that someone is worthy of blame—loses much of its original appeal. For if you and each of the other altruists were entirely isolated, then from your perspective, the situation would be identical to the one you face in 9,999 Mechanisms: with respect to the good outcome, there is nothing you can do but hope. In this case, as I see it, neither you nor anyone else seems all that blameworthy for failing to contribute. If others share my intuition here, it suggests that the possibility of cooperation plays an important role in our intuition about Drops of Water.

Putting the preceding points together: it seems plausible that the attitude of the altruists that we take issue with must concern cooperating with others. In particular, I think, this attitude concerns a *willingness* to cooperate—or to act in concert with others—where this cooperation has as its aim producing the good outcome. Since a willingness to act is, I think, no more than an intention to act under certain circumstances, or a conditional intention to act, it seems the relevant attitude here is a conditional intention to add one's water to the cart. Conditional on what? Well, given the aim of cooperating is to produce the good outcome, it seems your intention must be conditional on whether the collective set of acts would produce this outcome. Thus the intention must be to add one's water if and only if enough other potential contributors so intend—enough, that is, to ensure that the suffering of those in the desert will be relieved.¹⁶

¹⁶ Here I have in mind something similar to what Michael Bratman (2014, p. 72) calls a shared intention involving obligation-based interdependence.

Three things are immediately worth clarifying here. First, the content of this intention is meant to be understood from the perspective of the intender. So while the intention is to contribute if and only if enough others so intend, this does not mean that it must in fact be the case that enough others so intend in order for one to act on this intention. To act on the intention, one must only believe that enough others have an analogous intention. If one believes this, then one takes it as true, or represents the world as its being the case. And, from one's own perspective at least, that is as good as its actually being true. Second, and relatedly, it is worth pointing out the significance of the proposed intention's being conditional, on the present account, as opposed to unqualified or outright. A conditional intention becomes an outright intention when the intender forms the belief that the relevant condition has been met. Thus if one has the proposed conditional intention, one's forming the belief that enough others have an analogous intention will suffice to make one's intention outright—or, to make the content of one's intention just *to contribute*. Finally, and crucially, where the proposed intention concerns the intentions of “enough” others, what is meant is not just that there is a sufficient number of others, whoever they may be, who each possess an analogous intention, but is instead that there are enough other particular agents—such as Altruist 1, and Altruist 2, and Altruist 4, and so on—who each possess an analogous intention. So the idea is that each potential contributor's intention will concern the particular, individual intentions of each of the other potential contributors, where these are collectively considered to be either enough or not.¹⁷

¹⁷ This means that, if one were merely informed that “enough others” (whoever they may be) each possess an analogous intention, one may have no reason to form the conditional intention, or to contribute, on the current approach. This is because one's own intention would not concern the intentions of any particular others, and none of the particular others' intentions would concern one's own intention. While this conclusion may initially seem troubling, I think it actually is exactly right. Since one would believe that enough others intended to contribute, and thus (we can assume) that the sufferers would receive enough water, it is difficult to see how one could be rationally compelled to act so as to make the state of affairs in which the sufferers would receive enough water more likely. After all, anything one did would, by one's own lights, make no difference to the desired outcome—the sufferers would receive enough water in any case. What is needed is instead for one's own intention and the intentions of others to depend on each other for their motivational force, so that one's belief that enough others have formed the relevant intention cannot lead one to conclude that one's own intention or contribution would make no difference.

To be a bit more precise, then, the suggestion here is that you ought to form the following conditional intention in Drops of Water: to add your water if and only if (you believe that) enough other particular altruists each form an analogous intention. This, I claim, is the “something more” we each could do.

Now, I must admit: all of this remains, for the most part, rather speculative. To see that it is on the right track, though, consider a final variant of Drops of Water:

Slips on Water. Your situation is just as it is in Drops of Water, except now the altruists are not altruistic. They do not care in the least whether the suffering people receive enough water, and they have no intention of contributing their pints to the cart. Happily, though, these people are standing on a very slippery floor, and there is a good chance that when each goes to announce that they will not be contributing, they will accidentally slip, fall, lose hold of their water, and propel it into the cart. You are just as confident that enough of these people will accidentally add their water to the cart in this way to ensure the relief of the people in the desert’s suffering as you are that, in Drops of Water, enough other altruists will add their water.

Again, intuitions here may diverge, but I personally see little difference between this case and 9,999 Mechanisms. Indeed, my intuitions about the two cases are nearly identical: the claim that you should contribute in 9,999 Mechanisms seems less plausible than the claim that you should contribute in Drops of Water to roughly the same degree that the claim that you should contribute in Slips on Water seems less plausible than the claim that you should contribute in Drops of Water. Yet if our intuitions have this structure, it suggests that the relevant difference between Drops of Water and all of the other cases we’ve considered is that only in Drops of Water is it a real

possibility that the other contributors intend to contribute. That is, only in *Drops of Water* might it be the case that the other contributors have any intention of contributing their water to the cart. Since, in the mechanism cases, the other potential contributors are incapable of intention, and in *Slips on Water*, they are unwilling to intend to add their water, these seem to be cases in which any intention on your part to add your water to the cart will surely do nothing to influence the outcome. And, they seem to be cases in which you intuitively have little reason to contribute. This is best explained, I think, by the suggestion that the condition that must be met in order for your conditional intention to motivate you to act has clearly not been met, and indeed cannot possibly be met, in these cases. That is, since it is stipulated that there are not enough other potential contributors who each intend to add their water to the cart, you of course do not believe that there are, and so do not see the condition on which your conditional intention to contribute rests as met. That is why it seems plausible that you have no reason to contribute in these other cases, yet it also seems plausible that you do have such a reason in *Drops of Water*.

I have suggested, at any rate, that the attitude we are looking for might make its possessors such that their act of adding water would make a difference to the good outcome. Such an attitude would then solve the problem we face in *Drops of Water*, since it would follow directly from the fact that one's act of adding water would make a difference that one had a reason to add it. Given that everyone else would also have access to this reason, it seems that as long as each person acted in accordance with their reasons, enough water would be added to relieve the suffering of those in the desert.

Does our conditional intention have the relevant property? It does. This is of course not to say that forming the conditional intention alone is enough to ensure that one's act of adding water would make a difference. That would mean that your act of adding water would make a difference, even if you were the only one who had formed the conditional intention. And that is not plausible.

Rather, my claim is that forming the conditional intention is enough to ensure that your act of adding water would make a difference, so long as you and enough particular others each truly believe that you have each formed an analogous intention. This is because, if you and these others have formed the conditional intention, then each of you intends to add your water if and only if the particular others so intend. And so, if all goes well—that is, if each of you forms the relevant belief about the others’ particular intentions and reasons as you should—each of you will act on your (then-outright) intention to add your water to the cart, and the sufferers will thus receive enough water. Importantly, notice that it will then be true of each of you that, had you acted differently, the suffering people would not have received enough water—that is, each of your acts made a difference. After all, if you had not added your water, each of the other contributors would not have added theirs: just as your intention to contribute was contingent on each of their intentions, so each of their intentions were contingent on yours.¹⁸ Hence if you had failed to contribute, you would at some point have failed to possess the relevant intention. In which case, the other contributors would have abandoned their belief that enough particular others each had an analogous intention—that is, the condition that was needed for their conditional intentions to have any motivational force would not have been met. Thus, each of you would have had no reason to add your water to the cart; and thus, so long as each of you would have acted in accordance with your reasons, the suffering of those in the desert would not have been relieved. Since a similar counterfactual story could be told about each of the other contributors, it follows that each of your acts of adding water made a difference. Accordingly, forming the relevant conditional intention is enough to ensure that your act

¹⁸ This is why it is important that your intention is not conditional merely on whether “enough others” (whoever they may be) form an analogous intention, and is instead conditional on whether it is true of enough other particular agents that they form an analogous intention. The first approach would fail to make the intentions depend on each other in a way that made their motivational force depend on each of the others’ particular intentions. On that approach, it would only matter that “enough” others had the intention, in which case the loss of one intender (and so you in particular) would presumably not affect whether that condition was met. The current approach, on the other hand, makes each individual intention interlock with, and depend for its motivational force on, each of the other particular intentions, which means that the loss of any one intender would lead to the loss of each of the others.

of adding water would make a difference, so long as you and enough particular others each truly believe that you have each formed an analogous intention.¹⁹

That is all well and good, it may be thought, but this account ultimately fares no better than the ones considered earlier. Even if forming the relevant conditional intention would give us a reason to add our water, and would thus lead to the good outcome, it doesn't follow that we ought to form this intention. Much like the views we considered earlier, the current account offers a reason for us to contribute, but no plausible explanation of why we would be obligated to do so. This explanation cannot be, say, that forming the conditional intention would give us a reason to add our water, which would be a good thing to do, because we know that our act of adding water would make no difference to the good outcome. Accordingly, short of a more plausible explanation as to why we are obligated to form the conditional intention in *Drops of Water*, the current account will seem no more plausible than views that derive a reason to contribute from an ungrounded obligation to do so.

¹⁹ It is also worth mentioning that this solution to the problem of collective impact can be naturally adapted to apply to the analogous, intrapersonal problem of rational self-torture (Quinn 1990). The latter problem concerns a series of choices faced by a subject who is hooked up to a machine that can apply electric current to the subject's body in unnoticeably small increments. The machine has settings from 0, where the machine is off, to 1,000, which would be torturous. Each week the subject must choose: advance the machine's dial one setting—where this can never be reversed—and receive \$10,000, or stay put and receive nothing. Since each step between settings entails an unnoticeable difference in pain, there seems to be little reason not to advance to the next setting each week. And since receiving \$10,000 would be good, there seems to be good reason to advance to the next setting each week. If this reasoning is applied each week, however, the subject will eventually be in torturous pain, which they will gladly return their acquired fortune to relieve. The account I've provided here suggests the following solution. At the start, the subject should choose a setting, call it S_p , whose application leaves them in a state that they prefer to their current one (even if S_p is 1), and then intend to advance to the next setting each week until S_p is reached if and only if they intend to not advance to S_p+1 . Thus, they should form a conditional intention to increase the setting each week iff (they believe that) they intend not to advance to the next setting once S_p is reached—just as, in *Drops of Water*, the subjects should form a conditional intention to add their water iff (they believe that) enough particular others so intend. While the choice of any particular S_p may be in one sense arbitrary (since there will presumably be an S_p+1 whose application would involve no noticeably greater pain yet noticeably more money), the subject here at least does better than the one who is forced to choose between torturous pain and passing up on an effectively free \$10,000. For a similar approach to this problem, see Portmore 2019.

So why should we form the conditional intention to add our water if and only if enough particular others so intend? Here is an argument, in simple premise-conclusion form, as to why you ought to form this intention:

- P1 You should desire that the suffering of the people in the desert is relieved.
- P2 So long as your desire, if you have one, to keep hold of your water is weak, your desire that the suffering of the people in the desert is relieved should be stronger than your desire to keep hold of your water.
- P3 So, if you are rational, you should prefer the state of affairs where your options are
- (1) relieve the suffering of the people in the desert, or
 - (2) keep hold of your water
- to the state of affairs where your options are
- (1*) make no difference to the suffering of the people in the desert, or
 - (2*) keep hold of your water.
- You should prefer the state of affairs in which adding your water would make a difference with respect to the suffering of the people in the desert to the state of affairs in which adding your water would not make such a difference. That is, you should desire to become such that your act of adding water determines whether the suffering of the people in the desert is relieved.
- P4 Forming the conditional intention to add your water if and only if enough particular others so intend is a necessary means to becoming such that your act determines whether the suffering of the people in the desert is relieved.
- C You should form the conditional intention.

If this conclusion is rejected, I will be interested to see which premise the objector takes issue with.

To avoid any weak spots, let me briefly say a bit more about each premise.

P1 and P2 are, I take it, obvious. This is a point where I can provide no further argument.

Still, I'm confident in claiming that no plausible argument can be provided for denying these claims.

Why accept P3? Again, it seems to me too plausible to deny. The more general claim in P3 is the following. If you prefer an outcome x to an outcome y , which you prefer to an outcome z , and bringing about x and z would involve the same act, you should prefer the state of affairs in which your options are to bring about x or to bring about y to the state of affairs in which your options are to bring about y or to bring about z . I do not see a plausible counterexample to this claim.

Finally, P4 follows from what I have argued in the paper thus far. Forming the conditional intention to add your water if and only if enough particular others so intend might make you such that your act of adding water would make a difference. To ensure that possession of the intention would give you this property, though, something more is needed. In particular, you and these others each must truly believe that you each possess an analogous intention. P4 follows directly from this. Thus if we are to reject P4, an argument must be given as to why the solution to Drops of Water that I provide here fails. Again, I do not see a plausible argument of this sort.

The conclusion then follows. Notice that it applies equally to each of the altruists in Drops of Water. This means that, if someone does not form the conditional intention, it must be that they have failed at one of the steps along the way. Depending on which step this is, this person is likely to be worthy of at least some blame.²⁰ So, if a suboptimal outcome is brought about, and the altruists' attitudes are transparent to each other, it must be the case that someone is worthy of blame. More precisely, it must be the case either that: (i) not enough people formed the intention; or (ii) someone

²⁰ Though this blameworthiness may not in each case be explained by a failure of the same kind. As I see it, if one fails to satisfy P1 or P2, any blameworthiness seems to derive from a moral failing; and if one fails to satisfy any of P3-C, any blameworthiness seems to derive from either a rational or epistemic failing.

who formed the intention failed to act on it. As we have seen, in the latter case, the person is clearly worthy of blame: their act determined that the suffering people did not receive enough water. We can now see that, in the former case too, someone is worthy of blame. In particular, each of the altruists who were not willing to add their water if and only if enough particular others were so willing are blameworthy. And this is because they each must have failed at one of steps from P1-C.

This is a welcome result. It means that in any case where intuition suggests that someone is worthy of blame, the current account will be able to identify at least one such person. That's because, as we have seen, in cases where the altruists' attitudes are not transparent to each other, or where they cannot cooperate, no one seems all that blameworthy for a suboptimal outcome being brought about. It is instead cases where cooperation is possible that are intuitively the most troubling. And in these sorts of cases, if a suboptimal outcome is reached, the account I have proposed will locate at least one person who is worthy of blame. Thus if intuition is any guide, the current view must be on the right track.

Yet even if the view does get us the right results in these particularly troubling cases, still it might be objected that it fails to get us the right results in certain cases where the *good* outcome is reached. To see this, suppose that you are the very last altruist who must decide whether to add their water in Drops of Water. Since you and the other potential contributors are all altruistic, you've all formed the conditional intention. And, until you, everyone has contributed. But you might wonder why *you* should bother to contribute. Indeed, you might wonder why the altruist next to you bothered to contribute. And, it seems, you'd be right to wonder these things: after all, you're sure that enough water has already been added to the cart. The current proposal, however, seems to tell you to contribute. This is because you've formed the conditional intention, you believe that enough particular others have each formed an analogous intention, and you reason as you should. Worse still, the current account seems to imply that, if you contribute, then your act makes a difference. But

if you're sure the sufferers will receive enough water in any case, it seems implausible that your contribution would make any difference. So how *doesn't* the proposed view imply that you have a reason to contribute, and that your contribution would make a difference, in this case?²¹

The answer is revealed in our argument to the conclusion that you normally do have a reason to contribute. While you have formed the belief that would normally lead you to become motivated to contribute—or, the belief that enough particular others each have a relevant conditional intention—in this case you are not so motivated. This is explained, on the current account, by the fact that you no longer conditionally intend to contribute. You no longer have this conditional intention because the argument from P1-C is no longer true of you. Specifically, the argument breaks down at P4: it is no longer the case that your possessing the conditional intention is a necessary means to your being such that your act determines whether the suffering of the people in the desert is relieved. That's because, since it is already determined (at least from your perspective) whether their suffering will be relieved, it is no longer the case that this is something you can determine. Since you're aware of this, you are no longer rationally compelled to possess the conditional intention, and so you abandon it.

So I don't think, and the account does not imply, that you are rationally compelled to contribute when you believe that the sufferers will receive enough water. Ultimately, this is because you believe that your desire for the suffering of those in the desert to be relieved is (or, at least, will be) satisfied, which means that there is nothing, from your perspective, that can be done to make its satisfaction more likely. In cases of this sort, you are no longer rationally compelled to act in ways that would make the satisfaction of your desire more likely. And so, you are no longer rationally compelled to possess the conditional intention. Like the reason to form the conditional intention in the normal case, the lack of reason to form it here generalizes: so long as one believes that the desire

²¹ Thanks to an anonymous reviewer for pressing me on this.

that motivates one to form the conditional intention is (or, at least, will be) satisfied, one is no longer rationally compelled to form the conditional intention. Since this is true of each altruist, each altruist can recognize this belief as being an acceptable reason for another's failing to contribute (at least so long as the belief is formed in an acceptable way). So, if one reasonably believes that one's desire in P1 is satisfied, one is not rationally compelled either to form the conditional intention, or to contribute; and since one is not rationally compelled to form the intention, one's act of contributing here could not be motivated by the conditional intention linking one's own act to each of the other altruists' acts, which means that one's own act would not make a difference, on the current account. It seems, then, that if you believe that your desire for the suffering of those in the desert to be relieved is satisfied, you are not rationally compelled to form the conditional intention, and your contribution would not make a difference; yet, so long as you do not believe this, you are still rationally and morally compelled to form the conditional intention, and it still might be that your act would make a difference. This result seems to me exactly right.

Before closing, a final thought: what I've offered here is a way of reaching the optimal outcome in cases like *Drops of Water*. In real-world cases of this sort, however, we will not be fully informed as to the attitudes of others. We are unlikely to know whether enough others are willing, for example, to reduce their individual carbon footprint, or to vote for the preferred candidate in a national election. It will then not be enough, in these cases, merely to form the relevant conditional intention—something more will be needed in order to reach an optimal outcome. This means that, in cases where cooperation is possible but the attitudes of others are not transparent—or, any real-world collective action case—forming the relevant conditional intention will likely not be sufficient for avoiding blame. It might be, for example, that we are also required to communicate our intention to others, or to encourage others to form an analogous intention. Alternatively, it might be that we are required to protest or blame those who are unwilling to form the intention. What is demanded

of us will no doubt depend on the stakes involved—just how bad would the suboptimal president, atmospheric concentration of carbon dioxide, or relevant outcome be? But the point here is that, regardless of what *else* we are required to do, we will *also* be required to form the relevant conditional intention. If we are required to do more, it will be in the service of compelling others to form and act on an analogous intention. Thus we should think of the proposal here as, not a general solution, but a starting point for solving any collective action case.

4. Conclusion

In sum: there is indeed a previously unrecognized feature of non-contributors in cases like Drops of Water in virtue of which they might be blameworthy. In particular, this feature is a failure to possess a conditional intention to act in the relevant way if and only if enough particular others so intend. If each person reasons properly and desires as they should, each will form this intention. And, so long as cooperation is possible, the good outcome will be brought about. If a suboptimal outcome is reached, yet cooperation was possible, someone will be worthy of blame. This takes care of every relevant case where, previously, the presence of a wrong could be identified, but the source of the wrong could not.²²

²² Many thanks to an anonymous reviewer for truly helpful comments. And special thanks to Cheshire Calhoun and Doug Portmore for extremely helpful comments, discussions, encouragement, and advice.

References

- Barnett, Z. (2018). "No Free Lunch: The Significance of Tiny Contributions." *Analysis* 78: 3-13.
- Björnsson, G. and B. Brülde. (2017). "Normative Responsibilities: Structure and Sources." In K. Hens, D. Horstkötter, and D. Cutas (eds.), *Parental Responsibility in the Context of Neuroscience and Genetics*. Dordrecht: Springer.
- Bratman, M. (2014). *Shared Agency*. Oxford: Oxford University Press.
- Budolfson, M. B. (2019). "The Inefficacy Objection to Consequentialism and the Problem with the Expected Consequences Response." *Philosophical Studies* 176: 1711-1724.
- Cullity, G. (2000). "Pooled Beneficence." In M. Almeida (ed.), *Imperceptible Harms and Benefits*. Dordrecht: Kluwer Academic.
- Fanciullo, J. (2019). "What Is the Point of Helping?" *Philosophical Studies*. doi 10.1007/s11098-019-01263-7.
- Hill, Jr., T. E. (1979). "Symbolic Protest and Calculated Silence." *Philosophy & Public Affairs* 9: 83-102.
- . (1983). "Ideals of Human Excellence and Preserving Natural Environments." *Environmental Ethics* 5: 211-224.
- Jamieson, D. (2007). "When Utilitarians Should Be Virtue Ethicists." *Utilitas* 19: 160-183.
- Kagan, S. (2011). "Do I Make a Difference?" *Philosophy & Public Affairs* 39: 105-141.
- Nefsky, J. (2011). "Consequentialism and the Problem of Collective Harm." *Philosophy & Public Affairs* 39: 364-395.
- . (2015). "Fairness, Participation, and the Real Problem of Collective Harm." In M. Timmons (ed.), *Oxford Studies in Normative Ethics, Vol. 5*. Oxford: Oxford University Press.
- . (2017). "How You Can Help, Without Making a Difference." *Philosophical Studies* 174: 2743-2767.
- Norcross, A. (2004). "Puppies, Pigs, and People: Eating Meat and Marginal Cases." *Philosophical*

Perspectives 18: 229-245.

Parfit, D. (1984). *Reasons and Persons*. Oxford: Oxford University Press.

Pinkert, F. (2015). "What If I Cannot Make a Difference (and Know It)." *Ethics* 125: 971-998.

Portmore, D. W. (2018). "Maximalism and Moral Harmony." *Philosophy and Phenomenological Research* 96: 318-341.

---. (2019). "What's a Rational Self-Torturer to Do?" Unpublished manuscript.

Quinn, W. S. (1990). "The Puzzle of the Self-Torturer." *Philosophical Studies* 59: 79-90.

Regan, D. (1980). *Utilitarianism and Co-operation*. New York: Oxford University Press.

Sinnott-Armstrong, W. (2005). "It's Not My Fault: Global Warming and Individual Moral Obligations." In W. Sinnott-Armstrong and R. Howarth (eds.), *Perspectives on Climate Change*. Elsevier.

Tänssjö, T. (1989). "The Morality of Collective Actions." *Philosophical Quarterly* 39: 221-228.