

## The All or Nothing Problem<sup>1</sup>

Joe Horton

Suppose that two children are about to be crushed by a collapsing building. You have three options: do nothing, save one child by allowing your arms to be crushed, or save both children by allowing your arms to be crushed.

Here are two very plausible claims about this case:

(1) It is morally permissible for you not to save the children.

(2) It is morally wrong for you to save only one child.

(1) is plausible because of the sacrifice that saving the children requires. (2) is plausible because saving both children requires no greater sacrifice than saving only one. But there may be a problem with accepting both of these claims. Suppose that you are a bad person, and you dislike one of the children. You are willing to save the other child, but you are not going to save both. (1) and (2) seem to imply:

(3) You ought to save neither child rather than save only one.<sup>2</sup>

---

<sup>1</sup> For helpful comments and discussion, I am grateful to Stephen Bero, Renee Jorgensen Bolinger, Brian Carey, Richard Child, Stephen Finlay, Shane Gronholz, Nathan Robert Howard, Nicola Kemp, Nicholas Laskowski, Caleb Perl, Alexander Sarch, Thomas Sinclair, Jonathan Spelman, Anthony Taylor, Aness Webster, Jonathan Wright, an audience at the 2016 Rocky Mountain Ethics Congress, and the editors and referees at *The Journal of Philosophy*. I am especially grateful to Alexander Dietz, Abelard Podgorski, Jonathan Quong, Jacob Ross, Mark Schroeder, and Ralph Wedgwood.

<sup>2</sup> Equivalently: if you are not going to save both children, you ought to save neither.

And that seems very counterintuitive. Surely the best moral view would not discourage you from saving the one child.

This is just one instance of a more general problem. There are many cases in which, by making some great sacrifice, you could bring about either a good outcome or an even better outcome. In some of these cases, it is very plausible both that it is permissible for you to bring about neither outcome and that it is wrong for you to bring about the less good outcome. But together, these claims seem to imply that you ought to bring about neither outcome rather than the less good outcome. And that seems very counterintuitive. We can call this *The All or Nothing Problem*.<sup>3</sup>

In this paper, I develop The All or Nothing Problem, suggest a solution, and then draw out some implications for theory and practice. In section one, I consider three responses to the problem and argue that each is unsatisfactory. In section two, I defend a different response. I argue that we can plausibly accept a principle according to which, very roughly, if we are willing to make a sacrifice, we ought to bring about the best outcome that we can permissibly bring about by making this sacrifice. In sections three and four, I show that this principle has two important implications. First, it implies that if we are willing to perform acts that would normally be supererogatory, we ought to perform these acts. This implication might seem counterintuitive, but its counterintuitiveness can be explained away, and it has independent explanatory power. Second, this principle implies that, often, if we are

---

<sup>3</sup> This problem has not been discussed in the literature. Derek Parfit and Victor Tadros consider cases that give rise to the problem, but they are interested in these cases for other reasons. See Parfit, "Future Generations: Further Problems," *Philosophy and Public Affairs*, XI, 2 (Spring 1982): 113–172, at p. 131; Parfit, *On What Matters: Volume Two* (Oxford: Oxford University Press, 2011), at p. 225; and Tadros, *The Ends of Harm: The Moral Foundations of Criminal Law* (Oxford: Oxford University Press, 2011), at pp. 161–162.

willing to give to charity, we ought to give to one of the charities that would use our gift to do the most good. Since most people assume that it is permissible to give to whichever charities we like, this implication has great practical significance.

### I. THREE UNSATISFACTORY RESPONSES

In this section, I consider three potential responses to The All or Nothing Problem and argue that each is unsatisfactory. For concreteness, I focus on the collapsing building case, but it should be clear how each response would generalise.

The first response is to embrace *Act Consequentialism*. According to this view, we ought always to do whatever makes things go impartially best.<sup>4</sup> Since it is impartially better for you to lose your arms than for the two children to die, this view implies that you ought to save the children. It therefore implies that (1), the claim that it is morally permissible for you not to save the children, is false. If we reject (1), accepting (2) does not commit us to (3). We can instead accept the following claim:

(4) You ought to save both children, but if you are not going to save both, you ought to do the next best thing, which is to save one. That is, you ought to save one child rather than save neither.

People who are already inclined to accept Act Consequentialism might be satisfied with this response. And they might take The All or Nothing Problem to lend support to Act Consequentialism. But like many people, I find Act Consequentialism

---

<sup>4</sup> More precisely, according to Act Consequentialism, an act is permissible if and only if there is no alternative act that would make things go impartially better.

implausibly demanding.<sup>5</sup> It is very hard to believe that morality could require us to make great sacrifices whenever doing so would make things go impartially best. For this reason, this response seems to me unsatisfactory.

The second response is to reject (2), the claim that it is morally wrong for you to save only one child. We could claim instead that, though it is not wrong for you to save only one child, saving only one child would reveal your bad character and make you blameworthy.<sup>6</sup> I suspect that some people will be tempted by this response, but I think that it would be very hard to defend. Since saving both children requires no greater sacrifice than saving only one, saving only one child would be like leaving the other to die when you could save her without any sacrifice, and that would clearly be wrong. Furthermore, it seems very plausible that if our acts are not justifiable to the people whom they affect, these acts are wrong, and clearly you could not justify saving only one child to the child whom you leave behind.<sup>7</sup>

The third response is to reject the inference from (1) and (2) to (3). Again, I think that this response would be very hard to defend. The relevant inference is an application of the following principle:

---

<sup>5</sup> For example, Bernard Williams, "A Critique of Utilitarianism," in J. J. C. Smart and Bernard Williams, *Utilitarianism: For and Against* (Cambridge: Cambridge University Press, 1998), pp. 77–150; Samuel Scheffler, *The Rejection of Consequentialism* (Oxford: Oxford University Press, 1994), chapter 1; and Tim Mulgan, *The Demands of Consequentialism* (Oxford: Oxford University Press, 2005), chapter 1.

<sup>6</sup> For support for the view that we can be blameworthy without having acted wrongly, see T. M. Scanlon, *Moral Dimensions: Permissibility, Meaning, Blame* (London: Harvard University Press, 2008), chapters 1 and 4.

<sup>7</sup> For a defence of the claim that, if our acts are not justifiable to the people whom they affect, these acts are wrong, see T. M. Scanlon, *What We Owe To Each Other* (London: Harvard University Press, 1998), chapter 5; and Derek Parfit, "Justifiability to Each Person," *Ratio*, XVI, 4 (December 2003): 368–390.

(5) If *A* is morally permissible and *B* is morally wrong, we ought to do *A* rather than *B*.

This principle is intuitively correct. And there are countless cases that seem to verify it. Suppose, for example, that it is permissible to say something nice, permissible to say nothing, and wrong to say something nasty. (5) implies that you ought to say nothing rather than say something nasty. And that seems the right result.

I conclude that each of these responses—rejecting (1) by embracing Act Consequentialism, rejecting (2), and rejecting the inference from (1) and (2) to (3)—is unsatisfactory.

## II. A BETTER RESPONSE

I have just argued that three potential responses to The All or Nothing Problem are unsatisfactory. In this section, I suggest a different response. I first apply the response to the collapsing building case, and then explain how it generalises.

I think that we can plausibly reject (1) in favour of the following claim:

(6) If you were not willing to save either child, it would be permissible for you not to save either, but since you are willing to save one, you ought to save both.

Since what we ought to do does not normally depend on what we are willing to do, this claim might seem counterintuitive. But we can defend it as follows. It is very

plausible that if our acts are not justifiable to the people whom they affect, these acts are wrong.<sup>8</sup> And plausibly, it matters not simply that our acts are justifiable to these people, but also that they are justifiable in a way to which we could reasonably appeal. If you were not willing to save either child, you could reasonably appeal to having to sacrifice your arms as a justification for not saving both. But because you are willing to save one child, you cannot reasonably appeal to this justification. Since there is no other adequate justification for not saving both children, there is no adequate justification to which you could reasonably appeal. So it follows, on this view, that because you are willing to save one child, you ought to save both.

If we reject (1) in favour of (6), accepting (2) does not commit us to (3). We can instead accept the following claim:

(7) Because you are willing to save one child, you ought to save both, but if you are not going to save both, you ought to do the next best thing, which is to save one. That is, you ought to save one child rather than save neither.

This claim seems very plausible.<sup>9</sup>

It might be objected that, if (6) were true, you could escape the obligation to save both children simply by making yourself unwilling to save the one child. That would be a

---

<sup>8</sup> See footnote 6.

<sup>9</sup> It might seem that this argument succeeds only when you are willing to save the one child. If you are not willing to save either child, (6) implies (1), so accepting (2) commits us to (3). But what we are here committed to is an implicitly conditional version of (3): *if you are not willing to save either child, you ought to save neither rather than save only one*. I think that we can plausibly accept this conditional version of (3) so long as we also claim that: *if you are willing to save one child, you ought to save one rather than save neither*. We can plausibly accept the combination of these claims because, taken together, they do not discourage anyone who is willing to save one child from doing so.

bad result, for a moral view that incentivises you to make yourself unwilling to save the one child seems just as deficient as a moral view that implies (3).<sup>10</sup>

I am sceptical that we have enough deliberative control over what we are willing to do for you to simply make yourself unwilling to save the one child, but even if we do, this would not give you a way out of the obligation to save both children. If the only reason that you are not willing to sacrifice your arms to save the one child is that being willing would make you obligated to save both, you cannot reasonably appeal to having to sacrifice your arms as a justification for not saving both. Since there is no other adequate justification for not saving both, even if you make yourself unwilling to save the one child, you will remain obligated to save both.

I should emphasise that, on the view that I am defending, being unwilling to save either child is not a justification for not saving the children, but rather a condition of not being obligated to save them. What would justify you in not saving the children is the sacrifice that saving them requires. Whether you are willing to save either child just determines whether it is reasonable for you to appeal to this justification.<sup>11</sup>

We can now generalise this response to the problem. If there is no adequate justification of our acts to which we could reasonably appeal, these acts are wrong. When bringing about an outcome *O* would require us to make a great sacrifice *S*, we can normally appeal to *S* as an adequate justification for not bringing about *O*. But if we are willing to make a sacrifice that is not significantly smaller than *S* to bring

---

<sup>10</sup> I am grateful to an anonymous reviewer for pressing me on this point.

<sup>11</sup> It might also be worth emphasising that being unwilling to save either child does not make it wrong for you to save both children. That is because being unwilling to save either child does not make it unreasonable for you to appeal to the good of saving the children as a justification for doing so.

about an outcome that is not significantly better than *O*, and we do not have adequate agent-relative reasons to favour this other outcome, we cannot reasonably appeal to *S* as a justification for not bringing about *O*.<sup>12</sup> So if there is no other adequate justification for not bringing about *O*, we ought to bring about *O*.

The italicised clause in the above argument is important. We sometimes have good personal reasons to favour an outcome that is impartially worse. These reasons are, in the standard jargon, *agent-relative*. To see the purpose of the italicised clause, suppose that by making some great sacrifice, you could save either your child or two other children. Though saving the two other children is impartially better, you have adequate agent-relative reasons to favour saving your child. Were it not for the italicised clause, the above argument might have the counterintuitive implication that, if you are willing to make the sacrifice to save your child, you cannot reasonably appeal to having to make the sacrifice as a justification for not saving the two other children. But with the italicised clause, the argument avoids this implication.

The upshot of this argument is the following principle, which we can call

*Optimific Altruism*: If the only adequate justification for not bringing about an outcome *O* is that it requires a sacrifice *S*, and we are willing to make a sacrifice that is not significantly smaller than *S* to bring about an outcome that is not significantly better than *O*, and we do not have adequate agent-relative reasons to favour this other outcome, we ought to bring about *O*.

---

<sup>12</sup> Note that this claim applies not only when the sacrifice that we are willing to make is the same size as *S*, but also when this sacrifice is not significantly smaller than *S*. (If you are willing to sacrifice your arms to save one child, you cannot reasonably appeal to having to make a slightly greater sacrifice, such as losing your arms and bumping your head, as a justification for not saving both children.)



This principle is a little wordy, but the rough idea is simple: if we are willing to make a sacrifice, unless we have adequate agent-relative reasons to bring about a suboptimal outcome, we ought to bring about the best outcome that we can permissibly bring about by making this sacrifice. By accepting this principle, we avoid The All or Nothing Problem.<sup>13</sup>

A moment ago, we considered a case in which, by making some great sacrifice, you could save either your child or two other children. In that case, because you have adequate agent-relative reasons to favour saving your child, Optimific Altruism allows that it is morally permissible for you to save your child. We should now consider a version of this case in which you have no special ties to any of the children. Optimific Altruism might seem to imply that, in this case, if you are willing to make the sacrifice to save the one child, you ought to instead make this sacrifice to save the two other children. And that might seem problematic, for some people hold that, when we can save either a smaller group of people or a larger group of different people, even when we have no special ties to any of these people, it is permissible for us to save the smaller group (or at least to choose a group by running a lottery).<sup>14</sup>

---

<sup>13</sup> It might seem that Optimific Altruism can lead to a kind of circularity. This principle implies that what we are morally required to do can depend on what we are willing to do. But what we are willing to do sometimes depends on what we believe about what we are morally required to do. To avoid this circularity, we need to restrict when we count as willing to make a sacrifice *S* to bring about an outcome *O*. We count as willing to make *S* to bring about *O*, in the sense relevant to Optimific Altruism, if and only if we are willing to make *S* regardless of whether we are morally required to make it. Thus, if whether we are willing to make *S* depends on whether we are morally required to make it, Optimific Altruism does not imply that we are morally required to make it. I am grateful to an anonymous reviewer for prompting me to address this concern.

<sup>14</sup> For a defence of this view, see John M. Taurek, "Should the Numbers Count?," *Philosophy and Public Affairs*, VI, 4 (Summer 1977): 293–316. Taurek is sometimes interpreted as holding the view that we ought to decide which group to save by running a lottery that gives each group an equal chance. For a reply to Taurek, see Derek Parfit, "Innumerate Ethics," *Philosophy and Public Affairs*, VII, 4 (Summer 1978): 285–301.

I think that, in these cases, we are morally required to save the larger group. But despite appearances, Optimific Altruism is compatible with both views. If we think that, in these cases, it is morally permissible to save the smaller group, we presumably think that saving the smaller group is an adequate justification for not saving the larger group. Optimific Altruism then permits us to make a sacrifice to save the smaller group even when we could instead make this sacrifice to save the larger group. If we instead think, as many people do, that we ought to save the larger group, we presumably think that saving the smaller group is not an adequate justification for not saving the larger group. Optimific Altruism then implies that, if we are willing to make a sacrifice to save the smaller group, and we could instead make this sacrifice to save the larger group, we ought to make this sacrifice to save the larger group.

### III. RETHINKING SUPEREROGATION

I have argued that we can avoid The All or Nothing Problem by accepting Optimific Altruism. In this section, I argue that this principle has important implications for how we should understand supererogatory acts.

Optimific Altruism implies that, if the only adequate justification for not bringing about an outcome is the sacrifice that it requires, and we are willing to make this sacrifice to bring about this outcome, we ought to bring about this outcome. This implication might seem problematic. Suppose that you sacrifice your arms to save the two children from the collapsing building. We would normally regard you as heroic and praiseworthy. But since you did sacrifice your arms to save the children, you must have been willing to sacrifice your arms to save them. So Optimific Altruism implies

that you were obligated to save them. But if you were obligated to save them, why should we regard you as heroic and praiseworthy? We do not normally regard people as heroic and praiseworthy simply for doing what they are obligated to do.

We can press this worry further. When people make great sacrifices to bring about very good outcomes, we normally say that their acts were *supererogatory*, meaning that they were morally praiseworthy but not morally obligatory.<sup>15</sup> But because people who perform these acts must have been willing to make the relevant sacrifices, Optimific Altruism implies that, for these people, these acts were obligatory. If these acts were obligatory, they were not, at least on the standard definition, supererogatory. So Optimific Altruism implies that no one ever performs supererogatory acts.

These implications are counterintuitive, but their counterintuitiveness can be greatly reduced. If you sacrifice your arms to save both children, you were willing to sacrifice your arms to save them, so Optimific Altruism implies that you were obligated to save them. But though you were obligated to save them because of your willingness to save them, you were not initially obligated to be willing to save them. And it is largely because you were willing to save them, when it would have been permissible for you to give priority to your own interests, that we should regard you as heroic and praiseworthy. So though Optimific Altruism does imply that, at least on the standard definition, no one ever performs supererogatory acts, it leaves space for us to acknowledge another important way in which people go beyond what is morally required of them. They go beyond what is morally required of them by being willing

---

<sup>15</sup> For a discussion of supererogation, see J. O. Urmson, "Saints and Heroes," in A. I. Melden, ed., *Essays in Moral Philosophy* (Seattle: University of Washington Press, 1958), pp. 198–216; and David Heyd, *Supererogation* (Cambridge: Cambridge University Press, 1982).

to make great sacrifices even when it is permissible for them to give priority to their own interests.

If we accept Optimific Altruism, we could choose to revise the standard account of supererogatory acts. Rather than saying that acts are supererogatory when they are morally praiseworthy but not morally obligatory, we could say that acts are supererogatory when they are morally praiseworthy but not morally obligatory unless we are willing to perform them. If we accept this revised account, Optimific Altruism is compatible with the claim that people sometimes perform supererogatory acts.

And this revised account has independent explanatory power. People who perform apparently supererogatory acts often insist that they were just doing their duty. If we accept the standard account of supererogatory acts, this phenomenon is difficult to explain. We would have to claim either that people who perform these acts are often mistaken about what morality requires, or that they are often insincere when reporting their moral beliefs.<sup>16</sup> Since we take these people to be moral heroes, neither explanation is attractive. But if we accept the revised account of supererogatory acts, we need not attribute any mistake or insincerity to these people. On this account, because these people were willing to perform these acts, they were obligated to perform them.<sup>17</sup>

---

<sup>16</sup> For a discussion of this issue, see Alfred Archer and Michael Ridge, “The Heroism Paradox: Another Paradox of Supererogation,” *Philosophical Studies*, CLXXII, 6 (June 2015): 1575–1592.

<sup>17</sup> An anonymous reviewer writes, “there is a danger of overstating the extent to which the revised account makes it possible to accommodate the thinking of people who perform supererogatory acts. It is true that such people often say that they were just doing their duty, and that that claim would not be mistaken on the revised account. But they rarely say, and presumably rarely think, that they were obligated to act as they did only because they were willing to do so. To the extent that they think their duty was independent of their willingness to act, the revised account would have to say that they are mistaken.” As the reviewer says, people who perform supererogatory acts do not say, and presumably do not think, that they were obligated to act only because they were willing to do so. But I suspect that,

#### IV. RETHINKING CHARITABLE GIVING

I have just argued that Optimific Altruism has important implications for how we should understand supererogatory acts. In this section, I argue that this principle also has important implications for how we should approach charitable giving.

Most people think that giving to charity is supererogatory, at least when we have already given a significant amount.<sup>18</sup> Most people also think that it is permissible for us to give to whichever charities we like. But charities differ significantly with respect to both the kinds of help they deliver and the efficiency with which they deliver it. A gift to some charities will do much more good than the same gift to others. So Optimific Altruism implies that, if we are willing to give to charity, unless we have adequate agent-relative reasons to give to a suboptimal charity, we ought to give to one of the charities that would use our gift to do the most good.

It might be objected that, because giving to charity is supererogatory, it cannot be wrong for us to give to suboptimal charities. But we have already seen that this inference fails. It is supererogatory to save the children from the collapsing building, but that does not make it permissible for you to save only one child.<sup>19</sup>

---

on reflection, these people would agree that other people in their position might not have been morally required to act, even if they are unable to explain why they were morally required to do something that others might not have been morally required to do. Optimific Altruism offers these people an explanation: they were willing to make the relevant sacrifice, whereas others might not have been.

<sup>18</sup> Peter Singer is a notable exception. He argues that we ought to give to charity until giving more would require sacrificing something of comparable moral importance to the suffering that giving would prevent. See his “Famine, Affluence, and Morality,” *Philosophy & Public Affairs*, I, 3 (Spring 1972): 229–243.

<sup>19</sup> It might now be objected that, because different charities tend to help different people, giving to suboptimal charities is not like saving only one child when you could instead save this child and another, but rather like saving a smaller group of people when you could instead save a larger group of different people. As I mentioned at the end of section two, some people think that, when we must

When we compare some charities, there may be no fact about whether a gift to one would do more good than the same gift to another.<sup>20</sup> This is most likely when the charities have very different aims. For example, there may be no fact about whether a gift to a charity that aims to conserve natural beauty would do more good than the same gift to a charity that aims to restore eyesight to people with curable blindness.

But while some charities may be incomparable in terms of the good that they do, other charities are not. This is clearest when the charities have very similar aims. For example, there are several charities that aim to restore eyesight to people with curable blindness, and there is good data, readily available, indicating which of these charities are most efficient.<sup>21</sup> Even when charities have specific aims that are less similar, they may share more general aims by which we can compare how much good they do. For example, charities that aim to restore eyesight to people with curable blindness and charities that aim to prevent the spread of malaria seem to share the more general aim of reducing human suffering, and available data is at least suggestive about which of these charities are most efficient with respect to this more general aim.

If some charities are incomparable in terms of the good that they do, there is no charity that is uniquely optimal. But there are clearly many charities that are suboptimal, in the sense that they do less good than other charities with which they

---

choose whether to save a smaller group of people or a larger group of different people, it is permissible to save the smaller group. If they are right, Optimific Altruism has much more limited implications for charitable giving. But many other people think that, in these cases, we ought to save the larger group. I assume that view in what follows.

<sup>20</sup> I am grateful to an anonymous reviewer for pressing me on this point.

<sup>21</sup> Some relevant data is available at GiveWell.org, GivingWhatWeCan.org, and TheLifeYouCanSave.org.

are comparable. When a charity is not suboptimal relative to any other, we can count it as one of the charities that would use our gift to do the most good.

There is much room for disagreement about when we have adequate agent-relative reasons to give to suboptimal charities. It is likely that we have adequate agent-relative reasons to give to suboptimal charities that would significantly help our family, friends, or the other people to whom we have special ties. And we may have adequate agent-relative reasons to give to suboptimal charities that support projects to which we are strongly attached, such as museums or local conservation projects.

But while people sometimes give to charities because they have a special connection to these charities, many people also give to charities simply because they want to help make the world a better place. Optimific Altruism implies that these people ought to try to give to the charities that would use their gift to do the most good. As I have said, there is good data, readily available, to help them.<sup>22</sup>

## V. SUMMARY

I began by introducing The All or Nothing Problem and arguing that three potential responses are unsatisfactory. In section two, I suggested a different response. I argued that we can plausibly accept a principle according to which, roughly, if we are willing to make a sacrifice, unless we have adequate agent-relative reasons to bring about a

---

<sup>22</sup> See footnote 20. For further discussion of effective charitable giving, see Peter Singer, *The Life You Can Save* (New York: Random House, 2010); Peter Singer, *The Most Good You Can Do* (London: Yale University Press, 2015); and William MacAskill, *Doing Good Better* (New York: Avery, 2016). For a more sceptical assessment of our ability to compare how much good different charities do, see the symposium on effective altruism in the July/August 2015 *Boston Review*; Thomas Nagel's review article, "Ways to Help," in the November 2015 *Times Literary Supplement*; and Amia Srinivasan's review article, "Stop the Robot Apocalypse," in the September 2015 *London Review of Books*.

suboptimal outcome, we ought to bring about the best outcome that we can permissibly bring about by making this sacrifice. In sections three and four, I argued that this principle has two important implications. First, it implies that if we are willing to perform supererogatory acts, we are morally required to perform them. Second, it implies that if we are willing to give to charity, unless we have adequate agent-relative reasons to give to a suboptimal charity, we ought to give to one of the charities that would use our gift to do the most good.