

Doing Good and Doing the Best

When Leona Helmsley, a rather unpleasant American plutocrat, died in 2008, the estimated value of the estate she left was between three and eight billion dollars. She bequeathed a tiny fraction of this to certain relatives, pointedly omitting others. She also left \$12 million to her dog, an amount that was subsequently reduced to \$2 million by the Trustees of the estate. The remainder of the fortune went to the Helmsley Charitable Trust. The mission statement of the Charitable Trust stipulated that the money should be used, first, for “purposes related to the provision of care for dogs,” and, second, for “such other charitable activities as the Trustee shall determine.” The mission statement had earlier listed a third aim – “provision of medical and health care for indigent people, with emphasis on providing care to children” – but Helmsley later deleted that clause.¹

After her death, I was briefly interviewed over the telephone for an article on the legal challenges to her bequest. I was asked whether Helmsley acted impermissibly in leaving her money to dogs rather than to people. My response was rather simplistic and naïve. I said that

to give even two million dollars to a single little dog is like setting the money on fire in front of a group of poor people. To bestow that amount of money is contemptuous of the poor, and that may be one reason she did it. But to give such a large sum of money to dogs generally is not frivolous. I think it shows some misplaced priorities, but many bequests do. In a world where there is starvation and poverty, you can say that it's wrong to give money to universities, or museums, or, worst of all, to divide it up for your children and heirs who are already rich. Welfare for dogs is better than more pampering of the rich. It may indicate misplaced priorities, but it not frivolous or silly. It [the bequest “for the care of dogs”] is disgraced by the context, but the two bequests should be separately evaluated.²

Among the thoughts I had when I made this comment was that a major activity of charities that care for dogs is finding homes for stray dogs that would otherwise either be “euthanized” or left to live by scavenging, often in a diseased or injured condition, only to be eventually killed beneath the wheels of a car. Preventing a vast number of dogs from suffering one or the other of these fates seemed to me a worthy use of this woman's money.

But of course Helmsley's billions could have been used instead to prevent human suffering and to save the lives of a very large number of people. So the relevant issue

was essentially comparative: not whether preventing canine suffering and saving the lives of dogs was a *good* use of the money, but whether using the money for those purposes was morally acceptable when it was possible to use it in ways that would have done much more good. The Helmsley bequest thus raises the more general question whether it can be morally permissible to donate money to one charitable cause when one could instead donate the money in a different way that would prevent more suffering or provide greater benefits.

Most people assume that the money Helmsley controlled at the time of her death was hers to dispose of as she wished. If she had really wanted, for example, to convert it to cash and burn it in a latter-day “bonfire of the vanities,” even outside a shelter for the homeless, that would, many people assume, have been her right. Although I cannot argue against it here, I think that this view is mistaken. Not only would she not have had a “liberty-right” or permission to destroy the money, but she would also not have had a “right to do wrong” – that is, a “claim-right” or right against interference if she had attempted to destroy it. Helmsley had legal rights to more resources than it could have been morally justifiable for her to own, possess, or control. Most of her money, in my view, was legally but not morally hers to dispose of. That which she was not morally entitled to retain, she was morally required to give away. But because she had no moral entitlement to that which she was required to give away, her wishes in the matter of its disposal were morally irrelevant. It is arguable that the portion of her wealth to which she had a legal but not moral title belonged to no one and ought to have been used in a way that would have done the most good, impartially considered. When there are resources to which no one has a claim, the default assumption is that they ought to be used to do the most good, taking into account that this may involve giving some priority to those who are worst off.

But some of the money that legally belonged to Helmsley at the time of her death was also morally hers to control. Let us assume that she was morally entitled to leave that money to wealthy relatives or even to gather it while she was on her deathbed and burn it in her fireplace. In that case it would have been genuinely supererogatory for her to give it instead to organizations that would use it to provide “care for dogs.” Some philosophers think, however, that even when it is supererogatory to give money to a charity, once one has decided to give a certain amount, one ought to give it in a way that would do more rather than less good.³ These philosophers acknowledge, of course, that there are limits to the amount of effort one must devote to determining which charity would be most effective, but they do insist that, given the information that one can be reasonably expected to acquire, one must give the money in the way that would do the most good.

This view is, however, rather puzzling. One may wonder how, if it is permissible not to give at all, it could be impermissible to give to a less effective rather than to a more effective charity. Most people believe that not only *whether* to give to charity but also *to which* charity to give is entirely discretionary. When it is morally permissible

not to give to charity at all, it is also permissible to give to whichever charity one prefers, for whatever reason. It is generally believed, for example, that it is perfectly permissible to give to a charity that supports efforts to discover a cure for a rare and not terribly debilitating disease rather than to one that seeks a cure for a common and devastating disease. Indeed, many believe that it is no less laudable to give to the first than to give to the second.

Yet there do seem to be instances in which it is supererogatory to benefit another – that is, it is permissible not to provide any benefit – but in which one is morally required to provide the greater benefit if one provides any benefit at all. Derek Parfit has presented one such case.

“Suppose,” he writes, “that I have three alternatives:

A: at some great cost to myself, saving a stranger’s right arm;

B: doing nothing;

C: at the same cost to myself, saving both the arms of this stranger.”⁴

Parfit claims is that, although the cost to the agent of either of the two acts of rescue may make it permissible for him to do neither, once he accepts the cost of saving one arm, it becomes impermissible for him not to save the other. That is, while it is permissible to save *neither* arm, it is not permissible to save *only one* arm.

One might think that instances of charitable giving, such as the Helmsley bequest, are relevantly like this case of Parfit’s. For both in Parfit’s case and in cases in which charitable giving is supererogatory, the agent can choose between doing less good and doing more good at an equivalent personal cost. And one might thus infer that, because the agent in Parfit’s case is morally required to bestow the greater benefit, the same must be true in cases of charitable giving.

But in fact Parfit’s case differs from ordinary instances of supererogatory charitable giving in at least two relevant respects. First, once the agent in Parfit’s case has ruled out the option of doing nothing, he can either confer only one benefit at great cost to himself or confer that *same* benefit to the same person *and* confer *another* equally great benefit at no further personal cost. While this agent had sufficient reason to provide neither benefit, he has *no* reason to provide only one rather than both. To prevent the loss of only one of the stranger’s arms would be *gratuitously* to allow the stranger to suffer the loss of an arm. And to allow a great harm to occur when one could prevent it at no cost to anyone is wrong.

By contrast, if Helmsley had chosen to leave her wealth to charities that would have produced greater good by benefiting persons rather than dogs, she would not have produced the same benefits she in fact produced together with others. Rather, if she had given to more effective charities, that would have been worse for the beneficiaries of her actual bequest. For Helmsley’s action to have been relevantly like that of the agent

in Parfit's case, she would have to had to benefit only dogs rather than providing the same benefits to the same dogs while also benefiting people at no additional cost. In general, when one gives to a less effective charity, one does not gratuitously fail to benefit those who would have benefited, and by more, if one had given to a more effective charity. Each option would have victims in the sense that it would have been worse for those who would have benefited from the other.

The second way in which Parfit's case differs from ordinary charitable giving is that, as the case is presented, there is no reason why the agent might prefer to save only one arm rather than both, whereas people often have reasons for preferring to give to a less effective rather than a more effective charity. Helmsley, for example, like many good people who volunteer to work at animal shelters, cared about dogs. It was important to her to provide care for dogs that would otherwise have been killed or suffered miserable lives. There would therefore have been a cost to her in forgoing the option of helping dogs. Because of this, the cost to her of using her money to save people would have been greater than the cost to her of using that same money to save dogs.

In summary, in Parfit's case, doing more good rather than (or, more precisely, in addition to) less good has no cost for the agent and is better rather than worse for the beneficiary of the lesser benefit. But in the Helmsley case, giving a fixed sum to charities that would do more good would have been worse both for her and for those who benefited from her gift to charities that did less good.

The same is true of charitable giving generally. Most choices between preventing greater harm by giving to a more effective charity and preventing less harm by giving to a less effective charity are like Helmsley's choice rather than the choice that the agent in Parfit's case faces. People often give to a less effective charity because they have some personal reason for caring about the work of that particular charity, and their failure to prevent more harm is not gratuitous in the way that the failure to save the stranger's other arm is. This is because, if they had given to a more effective charity, that would have been worse for those who benefited from their donation to the less effective charity.

Another hypothetical example in the philosophical literature that may seem more closely parallel to the Helmsley bequest was presented some years ago by Shelly Kagan.

Suppose a building is on fire. Upon entering, I find a child and a bird trapped within. Needing one hand free to clear a path back outside, I can save only one of the two, and I hastily pick up – and escape with – the caged bird. Clearly I have done something wrong. Even if [a person with moderate views about doing good] believes that I was not morally required to risk my safety by entering the building in the first place, he nonetheless believes that once I have decided to undertake the risk, I

should have promoted the greater good, by saving the child. If my interests are equally affected by either of two courses of action, I have reason to pick that act with the objectively better outcome.⁵

This case is different from Parfit's in that, if the agent had chosen the option that would produce the greater good, he would not have achieved the lesser good and some additional good as well; rather, his saving the child would have excluded the saving of the bird. But the Helmsley case, in which she chose to save many animals rather than many people, may seem just like Kagan's case, only on a larger scale – that is, Kagan's case writ large. If that is right, and if Kagan's judgment about his own case is correct, we should conclude both that Helmsley ought to have donated her fortune to save people rather than to save dogs, and that, more generally, people who decide to engage in supererogatory charitable giving then acquire a conditional duty to give to the charity that, according to the evidence they can reasonably be expected to have gathered, would achieve the most good, or prevent or alleviate the most harm.

There is, however, one salient difference between the cases, which is that the agent in Kagan's case seems to save the bird on a mere whim. Just as in Parfit's case there seems to be no reason why the agent might prefer to save only one arm rather than both, so in Kagan's case there is no mention of a reason why the agent might prefer to rescue the bird rather than the child. Helmsley, by contrast, had a reason for saving dogs rather than people, which was that she loved dogs but apparently rather disliked most people. To make the cases relevantly parallel, we should imagine that Kagan's agent is a bird-loving misanthrope who enters the burning building with the aim of saving the bird. Because this agent cares enough about the bird to risk her life to save it, forgoing saving it would have been an additional cost to her (an opportunity cost) of saving the child.

Kagan concedes that this might make a difference morally. He writes that what his original "case suggests... is that although ordinary morality grants me the option to refrain from promoting the good in the pursuit of my interests, I do *not* have the option to react in a manner that neither promotes the good *nor* my interests."⁶ At least according to "ordinary morality," if the agent in Kagan's case has an interest in saving birds but not in saving people (so that it is better *for her* if she saves the bird), it might be permissible for her to save the bird rather than the child. Similarly, Helmsley's interest in saving dogs might have made it permissible for her to devote her fortune to the saving of dogs rather than to the saving of people. And, finally, people who engage in charitable giving often have an interest in the success of a particular charity – for example, it may matter to them to find a cure for a particular disease because someone they loved has died of that disease. In all these cases, then, one might claim that when it is supererogatory to do good at all, it can be permissible to act in a way that does less good if one has an interest in acting that way rather than in another way that would do more good. The explanation for this is that acting in the way that would do more good has an additional cost – namely, the sacrifice of the agent's interest in achieving the

lesser good. Because we are assuming that it is the cost to the agent of acting either way that makes her doing any good at all supererogatory, a decision to do *some* good cannot create a duty to produce the greater good if that would require a personal cost that is *even greater* than that which would be necessary to produce the lesser good, which itself is a cost that she is not morally required to accept.

Yet, in spite of this, Kagan is right to judge that the agent acts impermissibly in saving the bird. After she enters the burning building, she finds both the child and the bird. At that point, she has already taken the relevant risk. She has already made the personal sacrifice that made saving either potential victim supererogatory. Hence, she can no longer appeal to considerations of cost as a justification for not saving the child. Saving the child is thus no longer supererogatory. She is now in the position of someone who can save either a child or a bird, though not both, at no cost to herself (or at only a very small cost to herself). Such a person has a duty to save the child, thereby allowing the bird to die. It makes no difference if this person is a misanthropic bird lover who has a personal interest in saving the bird. That interest is comparatively minor. Saving the child rather than the bird would not be sufficiently costly to her to make it permissible for her to save the bird.

One might object to Kagan's claim that the agent's action is impermissible by reflecting on the case in the following way. Suppose that the agent is among many people who are gathered outside the burning building. Knowing that the risk involved in entering the building is sufficiently great to make it permissible for them not to conduct a rescue, they all decide not to enter – all, that is, except the one agent who dashes in and saves the bird. Can we really believe that, of all these people, the only one who acts wrongly is the only one who has done any good at all – and in conditions in which it was permissible for her not to do any? Although she has done less good than she could have, her failure to do more good is not gratuitous, as there were one or two individuals for whom her doing more good would have been worse – namely, the bird and perhaps herself, if she had an interest in saving the bird but not in saving the child.

This reasoning is, however, specious. As long as the agent remains outside the building with the others, her situation is the same as theirs. But once she is inside the building, her situation is relevantly different; for she, unlike the others, can save the child at no cost, and it therefore becomes her duty to save it. It is her failure to do so that is wrong. It makes no difference, moreover, how or why she entered the building. She would have had the same duty even if the others gathered outside had flung her in against her will.

This explanation of why the agent in Kagan's case must save the child rather than the bird is also the deeper explanation of why the agent in Parfit's case must save both the stranger's arms rather than only one. It is not in fact essential to his acquiring a duty to save both arms that saving both would involve producing the same lesser good and

an additional good as well. Rather, what is essential is that, in doing what is necessary to save one arm, the agent has already incurred the cost that made both options supererogatory. He has thus placed himself in a situation in which he can save a person's arm at no cost to himself or others. Assuming that he has a duty to save a stranger's arm if he can do that at no cost, he then has a duty to save the second arm. The situation would be much the same if, having incurred the cost necessary to save stranger A's arm, the agent found that he could either save A's one arm or save both of stranger B's arms, but could not save all three. I believe that he ought then to save both of B's arms, assuming that there is no relevant difference between A and B and that the loss of both arms is at least twice as bad as the loss of one arm. This is so even though, if he were instead to save A's one arm, he would not be *gratuitously* allowing B to lose both arms.

It does not follow, however, that Helmsley acted wrongly in failing to do what would have prevented the greater harm. When she was choosing whether to leave her money to charities that would care for dogs or to ones that would do more good by preventing the suffering and deaths of persons, she had not yet incurred the cost that made her giving to either type of charity supererogatory (which was mainly just the cost of giving her money away rather than spending it in a way that she might have supposed would posthumously benefit herself, such as having statues of herself erected at various Helmsley hotels). And the same is true in most cases in which people face a choice between giving to a less effective charity and giving to a more effective one. They do not incur the cost of giving prior to deciding to which charity to give; rather, they incur it when they give, so that their giving to any remains supererogatory. Their choice is thus unlike that which the agents in Parfit's and Kagan's cases face; that is, it is not between costlessly doing more good and costlessly doing less good. Only if their choice were of that sort would they be morally required to give to the most effective charity.

We can make a simple change to Kagan's case that makes it relevantly parallel to ordinary instances of charitable giving. In Kagan's original case, the child and the bird are in the same part of the burning building. When the agent enters the building, he has immediate access to both and can take either, but not both, with him on leaving. But suppose the building is large and that the child is near one entrance while the bird is near another. There is insufficient time for the agent to enter the building twice before it collapses. She must choose one of three options: go in one entrance and save the child, go in the other and save the bird, and not go in at all. The reason why it is wrong for the agent to save the bird in Kagan's original case does not apply in this version. In both versions, for the agent to be able to save the bird, she must have first incurred the risk of entering the building. In the original case, saving the child is still an option after she has incurred that risk. But in this further variant, if she incurs the risk necessary to save the bird, it is no longer possible for her to save the child. Suppose that she is standing outside the building deliberating about what to do. She knows that it is permissible for her to stay where she is, saving neither. But perhaps because she is a

bird lover, she dashes into the entrance near the bird and saves the bird, just before the building collapses. One cannot claim that she has acted impermissibly because she has failed to save the child when she could have done so at no cost, or only a very small cost, to herself. If she does wrong in saving the bird rather than the child, we still lack an explanation, or at least a non-consequentialist explanation, of why that is, given the assumption that it is permissible for her to rescue neither.

Kagan says of his case that “once I have decided to undertake the risk, I should have promoted the greater good, by saving the child.” But merely deciding to take the risk does not make it obligatory to save the child. It does not become obligatory to save the child until the risk has already been incurred, so that there is no further cost in saving the child. All this, of course, presupposes that saving either is supererogatory.

There are, as I mentioned earlier, many philosophers who believe, consistently with what Kagan says about his original case, that once one has decided to make a sacrifice by engaging in charitable giving, one ought then to make a reasonable effort to ensure that one does the most good by giving to the most effective charity. These philosophers will be disappointed by my argument to this point, which seems to exonerate Helmsley in her decision to use her wealth to save dogs rather than to save persons. Those who give their time to working at animal shelters rather than raising money to buy malaria nets may, by contrast, feel relieved. But perhaps both reactions are premature, as there are objections to my argument.

One is simply that it is difficult to believe that whether the agent in the burning building cases is morally required to save the child can depend on whether it is in the same room as the bird or in a different part of the building near a different entrance. That seems morally irrelevant. Yet, odd as it may seem, this is in fact relevant, as it determines whether the agent, having entered the building where the bird is, perhaps with the intention of saving the bird, can save the child at no risk or still faces the same risk in saving the child that she faced when she was outside. And, as I have repeatedly claimed, if she can save the child at no risk, she is morally required to save it, but if she can save it only at great personal risk, saving it is supererogatory.

One possible implication of my argument is that promoters of “effective altruism,” of whom I am one, might investigate whether the context of charitable giving could be arranged so that it would be relevantly like the situation of the agent who enters the building intending to save the bird but, once there, finds that she has a duty instead to save the child. But it is hard to see how one could permissibly rig the process of charitable giving so that those who, for whatever reason, want to give to a less effective charity could be led to make their donation only to find themselves morally obligated to ensure that it goes to a more effective charity instead.

To me it is counterintuitive to suppose that, in the revised version of Kagan’s case, it could be permissible for the agent to enter the building to save the bird when she

could, at no greater risk, go through a different entrance and save the child. But it also seems implausible to me to suppose that, in Parfit's case and the original version of Kagan's case, it is morally impermissible to produce the lesser good, given that it is permissible to produce no good at all. Suppose that in Parfit's case the agent is someone who is willing to save one of the stranger's arms at great personal cost but is, for some reason, simply unwilling to save both arms. It seems perverse to suppose that, given that morality does not require him to save both arms and that, in the absence of such a requirement, he *will not* save both, morality implies that this person may not save one of the stranger's arms when he is willing to do that, but must instead choose the permissible option of allowing the stranger to lose both arms.

Here is a rather more intelligible example that illustrates the same problem. Suppose that both a white person and a black person will die unless they are saved. The only person who can save them is a racist. This racist can

- A: at some great cost to himself, save the white person;
- B: save neither;
- C: at the same cost to himself, save both the white person and the black person.

Suppose that, because of the great personal cost of saving either person, option B is permissible. The racist is willing to accept this great cost to save the white person but is simply unwilling to save a black person. Indeed, the idea of saving a black person is so repugnant to him that he would rather allow a white person to die than save a black.

Assuming that morality does not require him to save both and that in the absence of that requirement he *will not* save both, my previous reasoning, which explains and justifies the judgments of Parfit and Kagan about this sort of case, implies that morality forbids the racist to save the white person despite his willingness to do it. I find that very hard to believe. It seems to me highly implausible to suppose that morality itself could require the innocent white person to pay with her life for the racist's attitudes.

One might argue that what morality actually requires is that the racist stop being a racist, in which case he would save both. It is certainly true that morality requires him to abandon his racist beliefs and attitudes. But that seems insufficient to solve the problem in this case if anything like the slogan "ought implies can" is true. For the racist cannot, in this emergency situation, instantly divest himself of the beliefs, attitudes, and habits of a lifetime, thereby becoming at least as willing to save both people as he is to save the white person.

A more promising argument might appeal to the claim that it is only the personal cost of the act of rescue – for example, the inevitability of injury or the risk of injury or death – that makes the act of saving supererogatory. If saving the black person were costless for the racist, morality would require him to do it. Indeed, even though the

saving would involve the comparatively minor cost of doing something the racist would find repugnant, morality still requires him to do it. It is only because any of the acts of saving would involve substantial personal cost to the racist that morality does not require him to save anyone, including the black person.

But if it is only this cost that exempts the racist from the duty he would otherwise have to save both, and if his reason not to save both has nothing to do with this cost, perhaps he is not exempted or excused from the duty after all. For he is willing to accept this cost as a condition of saving the white person and there is no further cost, apart from the repugnance, in saving the black person as well. Considerations of cost are, in effect, motivationally inert. Since the consideration that would release him from the duty to save both is in fact irrelevant to him where saving the black is concerned, he is left with the duty to save both.

Parallel claims apply in Parfit's case and in Kagan's original example. If these claims are correct, they make saving only one arm and saving the bird impermissible but do *not* make it impermissible to save one arm rather than save neither, or impermissible to save the bird rather than save no one; for they also rule out the option of doing no good at all. Whether this line of argument is correct is, however, an issue I will not pursue further here. For it has no application to ordinary charitable giving, which, unlike these cases, does not offer the option of doing the greater good at no further cost when one has already paid the cost of doing the lesser good.

This leaves me with the unwelcome knowledge that my arguments challenge rather than support some of the views associated with the effective altruism movement. What this suggests to me is that, at least in addition to exploring further arguments for the view that when one decides to do good, one ought to do the best, supporters of effective altruism should concentrate their efforts on showing that much less of doing good is supererogatory than we have hitherto imagined. It may be, for example, that saving both the white person and the black person is morally required at a much greater personal cost than we have previously supposed.⁷

¹ Jeffrey Toobin, "Rich Bitch: The Legal Battle Over Trust Funds for Pets," *The New Yorker*, September 29, 2008, p. 41.

² *Ibid.*, p. 47.

³ For example, Theron Pummer, "Whether and Where to Give," *Philosophy and Public Affairs* 44 (2016): 77-95.

⁴ “Future Generations: Further Problems,” *Philosophy and Public Affairs* 11 (1982): 113-172, p. 131.

⁵ Shelly Kagan, *The Limits of Morality* (Oxford: Clarendon Press, 1989), p.16.

⁶ *Ibid.*, p. 240.

⁷ I am greatly indebted to Sinan Dogramaci and Theron Pummer for highly illuminating comments on earlier drafts of this essay.