

Singer on Killing Animals

SHELLY KAGAN

There are, I think, at least two questions that any adequate account of the ethics of killing animals should try to answer. First of all, and most importantly, we'd like to know whether it is indeed wrong (other things being equal) to kill animals at all. Of course, killing may often involve pain, and most of us would agree that it is wrong (again, other things being equal) to cause an animal pain. But recognizing this fact doesn't yet tell us whether there is anything wrong with killing the animal *per se*—that is to say, above and beyond the pain it might involve. Suppose that we are considering killing a given animal painlessly. Would that still be objectionable? If so, why?

Second, assuming for the moment that it is, in fact, wrong to kill animals, is there something particularly wrong about killing *people*? That is to say, if we distinguish between being a *person* (being rational and self-conscious, aware of oneself as existing across time) and being what we might call a “mere” animal (sentient, but not a person), we might wonder whether it is somehow worse to kill a person than it is to kill an animal that is merely sentient. Most of us, I imagine, think that something like this is indeed the case.¹ But it is not obvious whether this common view is justified, and even if it is, it is not obvious what makes the killing of the person worse. (Of course, many people think it clearly wrong to kill people and not at all wrong to kill a mere animal; if that's right, then it follows trivially that killing a person is worse. But it is far from clear that this common view is correct.)

There is a third question—less central than these first two, but philosophically fascinating nonetheless—that merits examination as well. If

and when there would normally be something wrong about killing a given animal (whether a person or not), can that wrongness somehow be outweighed or canceled out by our creating a *new* animal, one that would not otherwise exist? If we *replace* the animal that is being killed, might that eliminate the objection to the killing that would otherwise be in order? Here, I suspect, most of us would want to put a tremendous amount of weight indeed on the distinction between persons and other animals. Many of us are at least tempted by the thought that merely sentient beings may be morally replaceable in just the way I have described. But it seems to be a different matter when it comes to people: people are *irreplaceable*. Or so most of us would want to insist. Once again, however, it is not at all clear whether this common view can be defended.

One of the most significant and influential discussions of all these matters can be found in Peter Singer's *Practical Ethics*.² Singer argues that it is indeed wrong to kill mere animals, though it is worse to kill people; and he defends the common view that people are irreplaceable, while mere animals are not. But the underlying philosophical issues are complicated, and in my own case, at least, I find that it isn't always easy to follow Singer as he tries to chart a course through the relevant philosophical thickets. Accordingly, in this essay I want to reconstruct the main lines of Singer's discussion, spelling out his view, and subjecting it to what I hope is friendly criticism. (To be somewhat more precise, I should say that what I want to examine is Singer's position as laid out in the third edition of *Practical Ethics*—for Singer readily admits to having changed his mind over the years, particularly with regard to our last question; and he remains uncertain about various aspects of that view.³)

KILLING MERE ANIMALS

Singer begins his discussion of killing by examining at length our second question—the question of whether there is anything especially wrong with killing a person. But I propose to start instead with a version of our first question, asking whether there is anything wrong with killing a sentient creature *per se*. Suppose that we are dealing with a mere animal, sentient but not a person. And suppose as well that we are contemplating killing it painlessly (a qualification I won't keep repeating). What, if anything, would be wrong with doing that?

A natural and plausible suggestion is this: killing an animal that would otherwise have an overall pleasant future prevents it from experiencing the pleasure it would otherwise get. That's why killing it is wrong.

Oddly, though, while Singer thinks something like this must be right (pp. 85–86), he also thinks that saying this requires us to adopt a view according to which pleasure has objective value. We can't give the plausible answer, he believes, if we restrict ourselves to the kind of preference view to which Singer is strongly inclined, a view according to which what matters is simply furthering the preferences of those affected (pp. 12–13). After all, a merely sentient being has no *present* concern for its *future* pleasure. So if we kill it, there is no preference that gets frustrated. Accordingly, Singer thinks, if we are going to say that it is wrong to eliminate the pleasure that would have come to the animal, we have to ascribe objective value to that pleasure.

But I don't see why the preference theorist cannot avail himself of the obvious suggestion despite all of this. While it may be true that a merely sentient creature has no current preference for the future pleasure, it seems plausible to suggest, nonetheless, that if the pleasure *had* occurred, the animal *would* have had a relevant preference—namely, for the pleasant experience to continue (as Singer notes, p. 86). So there would have been a preference, and that preference would have been satisfied, had we not killed the animal. If the satisfaction of preferences is good, why doesn't that suffice to ground the wrongness of killing the animal in question?

Singer doesn't explain himself, but my best guess is that he is here unwittingly presupposing the truth of something like what he later calls "the debit model" of preferences (pp. 113–14), according to which there is no positive value in the satisfaction of preferences *per se*; there is only disvalue in their frustration. On such a view, the mere fact that the animal would have a satisfied preference in the future (a preference for the pleasure it would then be experiencing) does nothing to introduce any positive value; and so preventing this satisfied preference from ever arising does nothing to reduce the overall balance of value from what it otherwise would have been. Consequently, if we are going to insist that killing the animal eliminates something of value, we have to embrace objective value after all.

If this *is* what Singer has in mind (and it is only a guess), then I want to make three quick remarks. First, Singer is apparently presupposing a view here—the debit model—which he hasn't even mentioned yet (it is only introduced almost thirty pages later). Second, and more importantly, the debit model itself has seriously unintuitive implications (which Singer himself is at pains to bring out, pp. 113–17), so preference theorists might be well advised to reject it in any event. And finally, Singer is mistaken, I believe, in thinking that he himself needs to embrace the debit model (a point I'll explain later); so even Singer might do well to avoid it. In which case, it seems to me, the preference theorist is indeed entitled to offer the

natural and plausible suggestion with which we started: killing a sentient creature is wrong (provided it will have an overall pleasant future) precisely because doing so prevents its future pleasure. As far as I can see, one simply needn't be an objectivist about value to say this.

As it happens, Singer also has a second worry about the natural suggestion (p. 87). If killing is wrong because of the loss of the potential future pleasure, by parity of reasoning don't we also have to say that it would be good to *create* extra sentient beings with pleasant lives? For that, too, would affect the total amount of pleasure (this time, in a positive direction). Yet many people find it implausible to hold that there is a significant moral reason to create extra sentient creatures merely because their lives would be pleasant ones.

This thought leads Singer to introduce some rival views concerning the ethics of creating new beings (pp. 88–90). He distinguishes between the “total view” (which accepts the implication in question, but which many find implausible for precisely this reason) and what he calls the “prior existence view” (which manages to avoid the implication, but which faces its own problems). But as far as I can see, Singer is mistaken in thinking that he needs to get into these issues at this point at all. For he is mistaken in thinking that the relevant counterpart to killing someone with a pleasant future is *creating* someone with such a future. Rather, it is *saving the life* of such a person.

After all, when we contemplate killing a sentient being, we are dealing with a being that *already* exists, not one that we will be bringing into existence. Accordingly, when asking what consistency commits us to here (with regard to aiding those with a pleasant future), we need only look at cases where the being we might aid already exists as well. Thus, if we insist—as we should—that it is bad to shorten a pleasant life (killing), all that consistency commits us to is the claim that it would be good to *lengthen* a pleasant life (saving). That is, if killing someone with a pleasant future is bad, then *saving* the life of someone with a pleasant future should be good. But this last judgment, of course, is one that virtually everyone will find intuitively acceptable.

As far as I can see, then, at this point in the discussion there is simply no need to get into the ethics of creating new beings at all. We need not ask whether it would be good to bring into existence a being with a pleasant future; we need only ask whether it would be good to *save* the life of a being who will then go on to have a pleasant future. And the answer, surely, is that this would indeed be good.

So the natural proposal stands. We can say that killing a merely sentient creature (with a pleasant future) is wrong because it robs the animal of the

pleasant future it would otherwise have. The preference theorist can say this just as well as the objectivist about value. Furthermore, offering the natural proposal doesn't yet require a foray into the ethics of creation.

But this is not to say, of course, that we can avoid the ethics of creation altogether. Singer is right to bring it in, although I think he brings it in at the wrong place. Where we need it, rather, is when we turn to the topic of replaceability.

Suppose our choice is this: we can continue to raise a happy, merely sentient animal, or we can kill this animal but also replace it, by creating a new happy animal (of the same kind) that would not otherwise exist. (Imagine that we cannot simply add the second animal while still raising the first; perhaps we lack the resources to raise more than one at a time.) On the face of it, it seems, if killing the first animal would normally be wrong by virtue of the future pleasure that thereby gets eliminated, then this wrong will itself be avoided (or compensated for) by the creation of the second happy animal, since this reintroduces the same amount of pleasure as the killing eliminated, leaving no net loss. Apparently, then, killing a mere animal is not wrong in cases involving this sort of replacement. Merely sentient beings are replaceable.

I take it that something like this is Singer's understanding of the basic line of thought that might lead us to the conclusion that mere animals are replaceable. Of course, as Singer hastens to point out, even if this thought is correct it won't actually support anything like current factory farming practices (p. 106); but in principle at least—if there is no flaw in the argument—it does seem as though a mere animal should be replaceable. Accordingly, it is at just this point that we need to ask whether it is really true that creating a new animal with a pleasant life counts as introducing a good that can be used to compensate for the bad done by killing the first animal and preventing its future pleasure. If not, then the argument for replaceability can be blocked. Here, then, the ethics of creation is directly relevant.

But there is a different objection to the replacement argument that should be considered as well, and as far as I can see Singer doesn't notice it. Even if the pleasant life of the new animal does count as a good, and as great a good as the potential future pleasure that is destroyed by killing, it doesn't yet follow that it is morally permissible to kill the first animal. To think that it does follow is to assume that moral permissibility is a simple matter of adding up the good and the bad consequences, so that an act is morally permissible provided that none of the alternatives would have better consequences overall. It is precisely because Singer does assume this—because he thinks that mere animals are appropriately subject to

a utilitarian calculus of this sort—that the permissibility of killing with replacement seems to be in play.

Note, however, that one might accept the initial thought that killing an animal is normally wrong because this prevents its future pleasure, and yet nonetheless reject the utilitarian treatment of animals, insisting, rather, that even mere animals should be handled within a *deontological* framework. Suppose, for example, that one were to hold that harm done to animals by robbing them of a pleasant future could *not* be justified by the mere fact that one will also bring about some further good. Then replaceability would not even be an issue: from a deontological point of view, the mere fact that an act's results may be good overall (or at least, not bad) simply won't suffice to justify doing harm.

Thus, those prepared to accord deontological rights to mere animals will not be moved by the possibility of replacement. And it does seem to me to be a shortcoming of Singer's discussion that he doesn't mention this sort of deontological approach here. Nonetheless, having noted this point I am going to put it aside, and I will follow Singer in thinking about these issues primarily from a utilitarian perspective.

Accordingly, we still need to ask whether it is really true that adding extra pleasant lives adds value. This is where the distinction between the total view and the prior existence view comes in. Singer notes, correctly, that both views have their counterintuitive implications (pp. 88–90). Indeed, I would want to add that pretty much every view here has some counterintuitive implication or the other.

Some jargon will be helpful. We're wondering what to say about cases where we face the possibility of creating an animal that would not otherwise exist. Call those animals who will exist *regardless* of the particular choice under examination the *inevitables* (since they will end up existing regardless of what choice I make now), and those animals whose existence depends on the particular choice being examined the *contingents*. Call those animals who actually end up existing at some time or the other—past, present, or future—the *actuals* (so all inevitables are actuals, but so are some contingents), and those animals that could have existed, but never do, *nonactual contingents*. If, for simplicity, we restrict our attention to the interests animals have in feeling pleasure and in avoiding pain, let us say that we *count* the pleasure a given animal would feel under some outcome if that gives us a reason to promote that outcome, while we count the pain they would feel if that gives us a reason to avoid that outcome. So our question becomes: Which interests are we to count? *Narrow* views count only the interests of the inevitables, while *wide* views count the interests of contingents as well; and

intermediate views count the interests of *some* contingents, but not all of them (for example, an intermediate view might count the interests of all actuals, contingent or inevitable, but not the interests of nonactual contingents).

Armed with these distinctions, here are some possible positions (there are more):

- The *total* view is a wide view. It counts the interests of everyone who will or might exist. And it does this symmetrically, taking into account both pleasure and pain.
- The *prior existence* view is narrow, only counting the interests of the inevitables. But with regard to those interests, it symmetrically counts both pleasure and pain.
- *Actualism* is an intermediate view, counting the interests of all and only actuals. Thus it counts the interests of all inevitables (since they must be actual) and of all actual contingents, but it does not count the interests of nonactual contingents. It too treats pleasure and pain symmetrically, when they count at all.

In addition to these various symmetrical views, there are also *asymmetrical* views, the most important of which are wide. These count the interests of both inevitables and contingents, but they do so in a way that treats pleasure and pain asymmetrically. Here are some versions of asymmetry (there are others):

- *Extreme asymmetry*, according to which pain counts against an outcome, but pleasure does not count in its favor—not even the pleasure of inevitables.
- *Moderate asymmetry*, according to which both pleasure and pain count for inevitables, but not for contingents. For contingents, only pain counts, not pleasure.
- *Impure asymmetry* (asymmetry with an “actualist twist”), according to which the interests of all *actuals* are treated symmetrically (that is, both pleasure and pain count for actual contingents, and not only for inevitables); but for nonactual contingents only pain counts.
- *Offsetting asymmetry*, according to which pleasure can count to *offset* pain (sufficient pleasure can cancel out the reason to avoid a given outcome that would otherwise be generated by the existence of pain under that outcome) but pleasure can never count robustly in *favor* of an outcome (there would, for example, be no reason to promote an outcome with pleasure but no pain).

- *Impure offsetting asymmetry* (offsetting asymmetry with an actualist twist), according to which pleasure *can* count robustly in favor of an outcome, but only in the case of actuals. For nonactual contingents, pleasure can only be counted to offset pain; it provides no reason in its own right to favor an outcome.

These are, I think, the eight most prominent views concerning the ethics of creation. Unfortunately, as I have already suggested, none of these views are easy to accept.

Initially, to be sure, actualism seems rather attractive: what could be more reasonable than to count the interests of exactly those beings who will exist at some time or the other? (Surely, it seems, we shouldn't count the interests of nonactual contingents—merely potential beings that never actually exist!) But this view leads to implausible paradoxes. Suppose for example that we must choose between outcome 1, where *A* will be the only sentient being that exists, at a mildly positive level of well-being (+10), and outcome 2, where *A* will be somewhat better off (+20), but *B* will exist as well, with a life containing far more pain than pleasure (−100). If I create outcome 2, then both *A* and *B* are actual, so the interests of both count, in which case it would have been better to create outcome 1 instead (since *A* would be only somewhat worse off, while *B* would avoid having the miserable life). Yet if I create outcome 1, then only *A* is actual, so *B*'s interests are irrelevant, in which case it would have been better to create outcome 2 instead (since *A* would be somewhat better off). So neither choice is acceptable! In effect, actualism ends up ranking outcomes in a way that implausibly depends on which choices we actually make.

Should we then accept the prior existence view? But this too gives an unacceptable answer in the case just discussed. When we face the choice between outcomes 1 and 2, *A* is the only inevitable, and so according to the prior existence view, *A* is the only being whose interests count. That means it is better to create outcome 2 (where *A* will be somewhat better off), even though this involves bringing into existence a sentient being—*B*—who will have a life of utter misery. According to the prior existence view, since *B* is a mere contingent her interests are irrelevant; but that seems unacceptable as well. (Singer makes a similar point with a similar case at p. 89; cf. pp. 108–11.)

The total view avoids this unacceptable implication, since it counts *B*'s interest in avoiding the life of utter misery. It insists, plausibly, that we must choose outcome 1 rather than 2. Even though it is true that, if we avoid creating her, *B* remains a nonactual contingent, her interests still count. However, the total view also implies that we have reason to create

extra sentient beings that will have pleasant lives. And as we have already noted, many find this implication implausible as well.

Asymmetry views can accommodate the intuition that there is no reason to create beings with pleasant lives, while still agreeing that there is reason to avoid creating beings with miserable lives. They do this, of course, by treating pleasure and pain asymmetrically. Roughly speaking, in the relevant cases they count the pain, but not the pleasure. But for exactly this reason such views strike most of us as unacceptably ad hoc (as Singer notes, p. 89). Absent an explanation for why pleasure should sometimes fail to generate a reason in cases where pain nonetheless would, asymmetry views seem philosophically unsatisfactory.

This vice—of being ad hoc—is common to all asymmetry views. But there are further specific objections to each particular asymmetry view as well. Extreme asymmetry has the unacceptable implication that there is no reason to make already existent sentient beings happier by increasing their pleasure. (Pleasure simply doesn't count, not even for inevitables.) Moderate asymmetry avoids this implication (since pleasures for inevitables do count), but it has the implausible implication that I act wrongly if I create a sentient being that has, overall, an incredibly pleasant life, as long as that being suffers any momentary pain whatsoever. (Recall that for moderate asymmetry, only the *pains* of contingents count, not their pleasures.)

Impure asymmetry avoids this last difficulty (since it counts the *pleasures* of contingent actuals, and not just their pains), but evaluations are once again implausibly dependent on one's choices (as with pure actualism). Admittedly, if I *do* create the being with the overall pleasant life, that is permissible (since the pleasure of the contingent actual counts and is more than sufficient to outweigh its pain). Yet if I do *not* create the being with the overall pleasant life, then—unacceptably—it turns out that it would have been *wrong* to create it! (For if I don't create this creature, it is a nonactual contingent; and for such beings pleasures don't count, so there is nothing to offset the pain, which does count.) Worse still, this view is doubly ad hoc: not only is it true that the *pain* of nonactual contingents counts, while the pleasure of nonactual contingents does *not* count, this very asymmetry only holds with regard to nonactual contingents. For actuals (whether contingents or inevitables), we have symmetry with regard to pleasure and pain instead.

From this perspective, offsetting asymmetry seems preferable. It avoids choice relative evaluation, and it says that it is permissible to create an animal with an overall pleasant life (since the pleasure offsets the pain). What's more, it isn't *doubly* ad hoc, since it treats pleasure as

a mere offsetter for pain in *all* cases (not just for actuals). Nonetheless, this view does remain implausibly ad hoc (since we have no explanation of why pleasure should count only insofar as it offsets pain, while pain is not similarly restricted). And it implausibly fails to recognize that we can have good reason to *increase* the pleasure of an already existing being—even if the animal already has an overall pleasant life. (On this view, one only has reason to increase the pleasure of existing animals up to the point where they reach the neutral level and all of the pain has been countered.)

Finally, while impure offsetting asymmetry avoids this last problem and recognizes that we have reason to increase the happiness of animals with lives that are already pleasant overall (since on this view pleasure counts robustly—and not merely as an offsetter for pain—in the case of actuals), it does this at the cost once more of being doubly ad hoc (like impure asymmetry): not only is it true, when it comes to nonactual contingents, that pleasure remains a mere offsetter (while pain is not similarly restricted), this asymmetry holds *only* with regard to nonactual contingents, and not at all for actuals. And of course—a last objection—the actualist twist behind impure offsetting asymmetry reintroduces the unacceptable dependence (already noted for actualism and impure asymmetry) of evaluations on one's actual choice.

As we can see, then, all the main views on this vexed topic are problematic. I've belabored this point because it should make it easier to recognize a potentially surprising fact, that the *correct* view—whatever it is—will inevitably have one or more features that will strike at least some of us as rather implausible.

So what *is* the correct view? Reasonable people, no doubt, can disagree about this; but for my money, at least, I am inclined to accept the *total* view. On balance, it seems to me, this is the least unacceptable of the various alternatives.⁴

If we do accept the total view—and if we keep in mind our earlier decision to follow Singer in thinking about mere animals within a utilitarian framework—then replaceability for mere animals does seem to follow. If I kill an animal with a pleasant life, I prevent the pleasure it otherwise would have had. But if I replace it with a new animal with an equally pleasant life, an animal that would not otherwise exist at all, then according to the total view this newly created pleasure counts as well, and by hypothesis it is as great a good as the pleasure that is prevented. So the result is no worse, overall, than had I simply allowed the original animal to live (and never created the replacement). Given a utilitarian framework, this means that my complex act—of killing and replacing—is morally permissible.

In short, given our assumptions, while killing a mere animal with an overall pleasant future is normally wrong, it will not be wrong if we replace that animal with another one. Merely sentient beings are morally replaceable.

KILLING PEOPLE

In the previous section I defended the claim that it is wrong to kill mere animals (with pleasant lives) on the ground that doing so robs them of the pleasure they would otherwise have. If sound, presumably this same line of thought applies to *people* as well. After all, nothing in that earlier argument turned on the idea that the animals in question were *merely* sentient beings, as opposed to being both sentient *and* rational and self-aware. So at a minimum, if some person would otherwise have a pleasant life, to kill her would normally be wrong, at least in part precisely because doing so would rob her of the pleasure she would otherwise have. Unsurprisingly, then, it is wrong to kill people, and not only mere animals.

But might there be something especially wrong with killing a person? Might this be wrong in some way that goes beyond the wrongness involved in killing any animal at all with a pleasant future? Singer thinks so, and mentions four possible considerations (pp. 76–85). All turn on the fact that since people are rational and self-conscious, aware of themselves as existing across time, someone with a pleasant future will typically have a desire that he continue to exist *into* the future. (Accordingly, these four considerations won't apply to the killing of mere animals.)

First of all, then, if I kill some person, this may well affect others who themselves have preferences about their continued existence (pp. 76–80). Killing one person creates anxiety (a kind of pain) in others, and this contributes to the wrongness of that killing. Now this first consideration certainly does seem relevant to the morality of killing; but since it is a highly indirect reason (and it depends as well on the possibility of being found out), Singer doesn't emphasize it.

The second consideration focuses on the preference of the victim himself (pp. 80–81). If—as is typically the case—the person wants to continue to exist, killing this person frustrates his preference. Indeed, it will normally frustrate many of the person's deepest and most central preferences. Therefore, if we accept the preference view—according to which it is wrong, other things being equal, to frustrate the satisfaction of a preference (and the stronger the preference, the greater the wrong)—then it will normally

be a great wrong to kill a person. It is this consideration, I think it fair to say, that Singer finds the most compelling.

But Singer mentions two other considerations as well. Third, then, it might be that people have a *right* to life, grounded in their desire to continue to exist into the future (pp. 81–83). Singer emphasizes the thought that having such a desire may be a necessary condition for having the right in question; so mere animals will lack this right. (Perhaps this explains why—as I noted earlier—Singer does not explore the possibility of treating mere animals within a deontological framework.) Finally, Singer notes that respect for autonomy may generate *further* reason not to kill a person, since killing a person typically interferes with their autonomous choice to go on living (pp. 83–84).

Although there are various subtle differences between the second consideration and the last two, all three can be seen as giving expression to the thought that if a person wants to continue living (if this is her preference, desire, or choice), then other things being equal this very fact makes it wrong to kill her. The most important place where they differ, it seems to me, is with regard to the *strength* of the reason in question not to kill. Singer understands the preference view in terms of utilitarianism, and as such the second consideration can, in principle, be outweighed, if conflicting preferences of others come into play. In contrast, the third and fourth views take the reason not to kill to be stronger—“more absolute,” as Singer puts it—less readily outweighed by “utilitarian calculation” (p. 81). In short, the third and fourth approaches express *deontological* perspectives (though Singer doesn’t use the term). They are intended to capture the view, held by many, that killing a person cannot be justified even when greater good might come of it.

This point is important, because it seems to me that Singer loses sight of it. Although he tells us that we should keep all four views “in mind” (p. 85), by the time Singer turns to a discussion of whether people are replaceable (as mere animals seem to be), the third and fourth views are forgotten; Singer considers the issue only from the perspective of the preference utilitarian. That discussion is certainly not without its interest. But it is a shortcoming nonetheless that Singer fails to so much as consider the issue from a deontological perspective as well.

Having noted this point, however, I am once again going to put it aside. Let us follow Singer and ask ourselves what a *preference utilitarian* should say about the replaceability of people. Normally, as we have seen, it is wrong to kill a person. But might it be permissible to kill someone *provided* that one also creates a *new* person (with an equally pleasant life) who would not otherwise exist?

Most of us, presumably, would insist that the answer is no. But Singer worries that the preference utilitarian may not be in a position to say this (p. 113). After all, while it is true that if we kill the first person his preference to remain alive is frustrated, the new person will *also* have a preference to be alive—and *that* preference *will* be satisfied. Since we end up with a satisfied preference either way, regardless of whether we kill and replace or simply let the first person be, the result seems to be no worse if we do kill and replace. So the preference utilitarian apparently has no ground for objecting. It seems, then, that according to preference utilitarianism not only are mere animals replaceable, people are replaceable too!

Singer takes this objection very seriously, and he significantly complicates his view so as to avoid the unwanted implication. One thing he does—as I have already noted—is to embrace the debit model of preference satisfaction (pp. 113–14). According to this view, recall, there is no positive value in the satisfaction of a preference, there is only negative value in the existence of unsatisfied or frustrated preferences. That is, although satisfying a preference eliminates or avoids the bad constituted by frustrated preferences (and so is good in this minimal, comparative sense), it introduces no robust, positive good in its own right. (Incidentally, although Singer explicitly endorses the debit model only with regard to *creating* preferences, as far as I can tell he means to accept it across the board, even with regard to already existing preferences.)

If the debit model is correct, then the argument for replaceability is blocked. When we kill the first person, this frustrates his desire to go on living, which is bad; and the fact that we also create a new person with a new preference—a preference that is satisfied—is simply irrelevant, since the satisfaction of that preference isn't good, but merely avoids a potential further bad (the bad that would occur if that new preference were unsatisfied). Thus, if I kill and replace, the result is actually worse (since the victim has a frustrated preference) than if I had merely let the first person be (for then his preference would be satisfied). In short, given the debit model, people are not replaceable.

Should we accept the debit model? Singer notes that it accommodates a certain number of our intuitions (pp. 113–14), but, as he also notes, this view also has an implication that most of us will find utterly unacceptable (p. 114): given the debit model and preference utilitarianism, it is wrong to have any children *at all*—no matter how happy they may be—since everyone has at least some unsatisfied preferences, however mild. After all, even the child's vast array of satisfied preferences can do no more than cancel out *most* of the “debit” thereby created, not all of it. So having a child must be wrong. (Similarly, on this view, a world with no sentient creatures

whatsoever would be better than an otherwise wonderful world in which anyone at all has *any* unsatisfied preferences; cf. p. 116.)

To avoid this latest implication, Singer complicates his view yet again (pp. 116–18). He entertains the possibility that pleasure is objectively good—not dependent for its value on being the object of preference. Indeed, he entertains the possibility that there may be other objective goods as well. If there *are* any such objective goods, then the objection can be answered: when we bring the overwhelmingly happy child into existence, the objective value of the pleasure, say, that she has (along with other objective goods, perhaps) outweighs the bad of having some unsatisfied preferences. So having the happy child remains permissible. (A similar answer, of course, applies to the example of the wonderful world.)

Unfortunately, Singer doesn't seem to notice that introducing objective goods into his theory reopens the door for the argument for replaceability. After all, if I kill and replace, it may be true that the satisfied preference of the new person doesn't count as a robust positive (given the debit model), but for all that, there will still be *objective* goods that are created as well (for example, the pleasure of the new person). Mightn't these outweigh the bad generated by frustrating the preference of the person we kill?

Since Singer doesn't discuss this worry, it isn't clear how he would reply to it. But perhaps we can suggest the following on Singer's behalf: Admittedly, when we kill and replace this will introduce some objective value, which may outweigh the bad due to the frustration of our victim's preference to live. But since, by hypothesis, the first person would have had a life as valuable as his replacement, it too would have contained objective goods (had he not been killed) of the same sort and significance. So the results are indeed better if we *avoid* killing: we have the same amount of objective value, while we avoid frustrating the preference of our potential victim.

Arguably, then, if he embraces both the debit model and the existence of objective value, Singer can indeed avoid the conclusion that people are replaceable. Nonetheless, it seems to me that there was never any need for Singer to introduce these various complications in the first place! For Singer erred, I believe, in thinking that preference utilitarianism straightforwardly implies replaceability. On the contrary, as far as I can see, given future-regarding preferences of the sort we are already entertaining, replacement will not normally involve as good a result as refraining from killing. Put simply, Singer failed to take into account some of the relevant preferences.

Here is a simple example to demonstrate the idea. First, let's adopt a fairly straightforward (if overly simple) version of preference utilitarianism: For each frustrated preference we assign a score of -1; and for each

satisfied preference we assign a score of +1. (Recall, we are looking to see if the preference theorist is forced to embrace the debit model, so we are here doing without it. Accordingly, satisfied preferences are assumed to have positive value.) Next, suppose (again, for simplicity) that the normal life of a person consists of four moments or stages. Imagine that at each such moment a person will normally have a preference to be experiencing pleasure at that moment. But more than this, at any given moment, and for each future moment (of the four), a person will normally have a preference to experience pleasure at that later moment as well. Then at the first moment, t_1 , the person will have four relevant preferences: he will have a preference to be experiencing pleasure at t_1 , a preference to be experiencing pleasure at t_2 , a preference to be experiencing pleasure at t_3 , and a preference to be experiencing pleasure at t_4 . In contrast, at the second moment, t_2 , the person will have only three relevant preferences (one each for t_2 , t_3 , and t_4). At the third moment, t_3 , the person will have two relevant preferences (one each for t_3 and t_4); and at the final moment, t_4 , the person will have one relevant preference (a preference with regard to t_4). Finally, suppose that, if left alone, the person will experience pleasure at each of his four moments.

Given all of this, it is easy to see that if he is left in peace, all of the person's preferences will be satisfied. And since there are ten of them in total (4 in the first moment, 3 in the second, 2 in the third, and 1 in the last), this results in an overall score, with regard to this life, of +10.

Now let us compare this score to the score we get over the *same four moments* if the first person is killed after t_2 , and then immediately replaced with someone new, who will have her life begin at t_3 . In this case, the first person has only three of his preferences satisfied (his preferences at t_1 with regard to t_1 and t_2 , and his preference at t_2 with regard to t_2), while four of his preferences are frustrated (his preferences at t_1 with regard to t_3 and t_4 , and his preferences at t_2 with regard to t_3 and t_4). So far, then, this gives us a score of $3 + -4 = -1$. (Obviously enough, since he is killed after t_2 he never comes to have any preferences at t_3 or t_4 .)

And what about the replacement? Let us suppose that she is allowed to live out a normal life of four moments. Still, if we are only interested in the value that arises during t_1 – t_4 , then we should only look at the preferences of the replacement that exist during t_3 and t_4 (disregarding those that exist later, during t_5 and t_6). There are seven such preferences, all satisfied, for a score of +7. Adding this to the score of -1 generated by the satisfied and frustrated preferences of our first person, this gives a total score for the killing and replacement scenario of $-1 + 7 = +6$ for the period in question (t_1 – t_4).

[150] *The Ethics of Killing Animals*

In contrast, as we have already noted, had we let the first person live out his natural lifespan, the total score during the same period would have been +10! In short, even according to the preference utilitarian, killing a person and replacing him leads to worse results overall, during the course of what would have been the first person's life.

This example is, of course, unrealistically simplified; but not, I think, in ways that affect the main point.⁵ Given the fact that people ordinarily have preferences with regard to their future, killing someone with a pleasant life and replacing him will normally have worse results—over the same initial period of time—from the standpoint of preference utilitarianism. So it is not true in any straightforward sense that under preference utilitarianism it is just as good to kill someone and replace them as it is to let them be. Unlike mere animals, people are not replaceable in that way.

To be sure, it might be objected that if we take into account *all* of the results—not just those that occur during what would have been the first person's lifetime, but also those that occur afterward—preference utilitarianism might well sometimes conclude that the results would be better if we did kill and replace. In principle, after all, the satisfied preferences that the replacement person will have *later* in life might well outweigh the frustrated preferences of the victim. (That won't be true, as it happens, in our particular example; but it certainly can be true if the victim is killed late enough in life, and the replacement lives for a long enough time afterward.)

But that still doesn't mean that the preference utilitarian thinks it permissible to kill and replace. For if we are going to extend the time frame of our comparison, and take into account the benefits from replacement due to the increase in the total amount of life that replacement makes possible (that is to say, take into account the gains from satisfied preferences that occur *after* t_4), then it is important to bear in mind that replacement will provide even *greater* overall gains if it is done without killing. That is, the results would be better still if, instead of killing first and then replacing, we create the new person only *after* the first person has lived out his life. Letting live and *then* replacing will normally have better results than *killing* and replacing. So even when we calculate over the extended time frame, simple killing and replacement won't normally be permissible.

In short, the preference utilitarian won't normally approve of killing a person, even *with* replacement. The results of killing will normally be significantly worse than the results we would have under alternative courses of action that don't involve killing at all.

Where then does that leave us? None of this should be taken to show that preference utilitarianism gives acceptable answers across the *entire* range of cases that involve killing (with or without replacement).⁶ But it

does show, I think, that it is a mistake to suggest that preference utilitarianism straightforwardly implies that people are replaceable. At a minimum, the question is far more complicated than one might initially have thought. But more than this, I am inclined to think that we are justified in accepting a considerably stronger conclusion: the preference utilitarian does not view people as replaceable at all.

If I am right, then perhaps Singer never needed to complicate his position in the various ways that he did. Or rather, somewhat more precisely, he didn't need to do it so as to avoid saying that people are *replaceable*. For all that we have seen so far, Singer could have remained a simple preference utilitarian after all.⁷

NOTES

1. However, many people would also want to grant human infants and the severely cognitively impaired a status similar to (or identical with) that of being a person, even though these humans are not in fact persons in the technical sense I have just characterized. Those sympathetic to this view sometimes appeal to the given human's *potential* to become a person, or to their *membership in a species* whose typical adult members are persons (or perhaps, in some cases, to their *having been* a person). Whether anything like this could suffice to earn a nonperson a moral status similar to that of someone who actually *is* a person, is a complicated and controversial question that, for simplicity, I want to put aside here. In what follows, therefore, assume that the "mere animals" I am discussing are ones that lack any such relevant connection (if such there be) to personhood.
2. Singer 2011. All parenthetical citations in this chapter are to this work.
3. Indeed, Singer appears to have changed his mind yet again about some of the ideas discussed below. (See *The Point of View of the Universe* [Lazari-Radek and Singer 2014], published after this essay was completed.) Regardless, the discussion in *Practical Ethics* remains significant, whether or not Singer still accepts all of the terms in which it is framed.
4. For a contrary view, see Višak 2013.
5. For example, we might prefer to think of time as continuous (rather than discrete), with individual preferences lasting more than one moment. But a preference theorist can then hold that the value of satisfying a preference (or the disvalue of frustrating it) is proportional to the length of time that it is held. This will generate the same results as the simple model discussed.
6. The argument I have given assumes that one's future-regarding desires are *indexical*, rather than qualitative: I want it to be the case that *I* am happy—not merely that some creature be happy, nor even that some creature that thinks he is Shelly Kagan be happy. Presumably, of course, this is normally a realistic assumption. But imagine that in some case I want only that someone with my various memories, beliefs, and goals (and so on) exist in the future. Then certain types of killing and replacement—perhaps involving the creation of an exact qualitative duplicate of me—may turn out to be acceptable after all. That seems to me to be the right position to take, but I won't pursue the question here.



7. This is not to say that I think that preference utilitarianism *is* the best position to take. I myself would rather embrace a more objectivist view about value. For example, while Singer is right to suggest (pp. 90–93, 103–4, 122) that the lives of persons have greater value than the lives of mere animals—particularly those persons whose lives are richly interwoven due to a wide range of cross-temporal desires covering significant stretches of time—I would want to explicitly add what Singer does not, that this is a form of *objective* value. Although Singer’s (hesitant) turn to objectivism may not be required for the reasons he thinks it is, I believe he is right to take it nonetheless.

