

Book Review

Fellow Creatures: Our Obligations to the Other Animals, by Christine M. Korsgaard. New York: Oxford, 2018. Pp. 271.

Christine Korsgaard is one of the best-known living Kantian ethicists, a long-time friend of animals (her book is dedicated to her cats), and an outspoken vegetarian. A Kantian ethicist who is also an animal advocate is inevitably asked whether she agrees with Kant that non-human animals lack intrinsic moral standing. Since 2004 or so, Korsgaard has given lectures and published papers indicating that she does *not* agree with Kant on this issue.

In fact, she says, ‘I end up agreeing with the utilitarians about which creatures have moral standing’ – namely, the sentient ones (p. xii). But it hasn’t been clear exactly *how* she would depart from Kant, and whether her view would be similar to that of the other famous Kantian ethicist in recent animal welfare discussions, Tom Regan. *Fellow Creatures* is the much-anticipated product of around fifteen years of reflection, and gives us her full position on these matters.

As expected, the position is not Kant’s. Kant thinks we have duties to treat non-human animals humanely when they’re alive, and to kill them painlessly whenever possible. But, notoriously, these are not duties *to* the animals themselves. Rather, they are duties to *other humans* who own or need them, and (preventively) to *ourselves*, since we might ‘demean’ ourselves and become desensitized to human suffering if we get used to treating non-human animals cruelly or capriciously (*Lectures on Ethics* 27:710; see also *Metaphysics of Morals* 6:241).

Korsgaard, by contrast, believes that non-human animals are *our fellow creatures*: they have direct moral claims on us just as our fellow human beings do, and we must act and legislate with their final ends in mind. This inclusion of all animals in the Kingdom of Ends is what makes hers a *broadly* Kantian rather than a consequentialist or virtue-theoretic position (even if there is a heavy measure of Aristotelianism sprinkled throughout). What she is grasping for here is the holy grail for Kantians (me included) who are also animal advocates: everyone wants to find a principled way – within the Kantian system – to expand the *Formula of Humanity* into what might be called the *Formula of Sentient Animality*.

The challenge, of course, is to find a plausible argument – and preferably a naturalistic one – that does the trick. I will focus on the centrepiece of

Korsgaard's book: what I call her 'Master Argument' according to which sentient animals are ends-in-themselves with membership in 'the moral community'. Along the way I will point out how her argument (if not her conclusion) differs from Regan's.

1. Assembling the pieces

Fellow Creatures is an exceedingly well-wrought piece of philosophy, neatly divided into four main parts and, within each of the chapters, into thematic sections, sub-sections and sub-sub-sections, in good *Tractatus* fashion. Its cover is one of the most strikingly attractive of any philosophy book I own.

It is also written in a strikingly honest way: Korsgaard openly struggles with questions about whether we could ever reasonably try to get predators to change their behaviour (or phase them out), about whether experimentation on animals is ever permissible, about when and how we should kill insects (even dust-mites!), and about whether we should have pets. The discussion of the last question is particularly intriguing and indeed wrenching, given that we know how much the author loves her cats.

The main goal of the book is to make the following Master Argument:

- (1) Each of us correctly takes the fact that some things are good or bad for us as a defeasible basis for making moral claims on others. (A banal example: it is good for me to walk into the restaurant right now, and I take this fact to support a *prima facie* moral claim against you blocking my way.)
- (2) There is no reason to think that such goods cannot be the basis of moral claims for 'every creature for whom things can be good or bad', including all the sentient animals.
- (3) Thus, every such creature 'has moral claims on us' (p. xi).

Making the full case takes quite a while, and we are often reminded in the first seven chapters that 'I haven't yet presented my argument'. That much throat clearing has the effect of building suspense as we approach the Master Argument in Chapter 8. Along the way there are lots of interesting side-discussions (in particular, a naturalistic account of the atemporal noumenal self) and a careful curation of Aristotle's and Kant's most valuable insights for Korsgaard's purposes.

Here are the latter, by way of summary:

- From Aristotle, Korsgaard takes the idea that there is a kind of goodness that is 'tethered' to the needs and desires of a sentient being – that is, a being 'who experiences her own functional condition in a valenced way, and pursues her own functional goods through action' (p. 22). She calls this 'goodness *for*'. Aristotle focuses on needs and desires that are grounded in our natures and well-functioning, a main example of which is the desire to *live*. If we have the desire to live as part of our natural well-functioning, then regardless of whether our

life is good or bad by some other measure, it is certainly *good for us*. To kill a sentient creature that clearly has such a desire (even painlessly in the Michael-Pollan-happy-meat kind of way) is thus to deprive it of something that is *good for it* (pp. 21-22).

- From Kant, Korsgaard takes the idea that we can have no knowledge of supersensibles, and so even if there is a transcendent source or measure of value, it is of no use to us. Without such knowledge ‘all we have to go on is that some things are certainly good-for or bad-for us. That then is the starting point from which we build up our system of values...’ (p. 145). But if there is no measure of value that isn’t tethered to someone or some species, then the question whether the good for human beings is more valuable or important than the good for other animals is ‘nearly incoherent’ (p. 15). What is *good for* one kind of creature can’t be measured against what is *good for* another, just as what is functionally good for a knife (being sharpened) can’t be measured against what is functionally good for a computer (being updated).
- From her previous self, Korsgaard takes the idea (defended at length in her 2009 book *Self-Constitution*) that we play an active role in the ‘constitution’ of our own selves as final ends – when, for instance, we take what we care about to establish moral claims on other people. Korsgaard ascribes this idea to Kant, but notes that he failed to see that ‘the self comes in degrees’. So even if the other animals cannot self-constitute in this high-level way, she says, that does not mean that ‘there is no functional unity to [their] minds and selves’ (pp. 34-35). Likewise, even if non-human animals cannot value themselves and their goods as ends-in-themselves, that does not mean that they and their goods cannot be valued by *us* as such.

There is another discussion in the seven-chapter build up to the Master Argument that is worth mentioning – it has to do with how Korsgaard thinks reason can ‘test’ maxims via universalization. Some Kantian ethicists reject the idea that Kant views the categorical imperative as a test that can be implemented in practice. But Korsgaard thinks that some version of the test will often work, at least for perfect duties:

[T]o ask whether you can will a maxim as a universal law is to ask whether you can will the universal practice of pursuing a certain end by means of a certain kind of act without undercutting the effectiveness of that kind of act for achieving that end. (pp. 126-127)

In general, Korsgaard says, the test works well for maxims involving act types that ‘are supported by conventions or practices’ — like falsely promising to pay back a loan (p. 127).

She acknowledges, however, that there are other kinds of maxims for which the test doesn’t work well: maxims involving ‘nature’ rather than convention.

Kant's prohibition on suicide, for instance, is notoriously difficult to motivate via the universalization test – here commentators typically turn to the Formula of Humanity rather than the Formula of Universal Law. But this means that maxims involving acts directed towards non-human animal natures 'have precisely the features that put Kant's universal law test under most strain', because in these cases we *cannot* turn to the Formula of Humanity to save the day (p. 129).

Korsgaard's big step here – and her big departure from Kant, if not from Regan — is to expand the Formula of Humanity so that it applies to all sentient animality:

[W]hen we consider the reason why Kant thinks we must claim the standing of 'end-in-itself' for ourselves, we will see that we must claim that standing for the other animals as well. (p. 130)

On her view, other animals are our *fellows* in the moral community, and not just our tools and property: 'That is what the "fellow" in "fellow creatures" does – it gestures at a relation in which we stand to each other' (p. 96).

The most prominent animal advocate in the Kantian tradition up to now, Tom Regan, argued for a similar conclusion: sentient creatures are 'subjects of a conscious life', just as we are, and this counts as strong evidence that they too possess the 'inherent value' that is the basis of moral standing and rights. Regan leaves the latter notion as a value primitive: there can be no account of what *inherent value* consists in. He also insists that being a subject of a conscious life is just an indication, and not a necessary condition, of having this kind of value (thus someone in a vegetative state who is not a subject of a conscious life still has inherent value).

Korsgaard eschews such appeals to a mysterious kind of value known in a mysterious kind of way, even if her ultimate conclusion (that all sentient animals are ends-in-themselves) is the same as Regan's. Motivating this big step away from Kant in a *naturalistic* way is the task of her Master Argument.

2. The Master Argument: an overview

Chapter 8 is the main event. Here Korsgaard gathers the pieces she has collected – from Kant and Aristotle, and from her own independent argumentation — and masterfully combines them into 'a Kantian case for our obligations to the other animals' (p. 131).

As we have seen, she starts off by supposing that there are no intrinsic values (or that, if there are, we are unable to know about them). All we have to go on is the way we finite valuers value things – that is, the way in which things are *good for* us, given our desires and preferences. Other naturalistic philosophers and economists leave the individual story there and move on to a view about how we can coordinate these various preferences. But Korsgaard thinks that two further things are worth noting at this stage. She speaks of them as the two logical 'moments' involved in 'making a choice' (p. 147).

First, as we have seen, the fact that I desire some outcome, or that my nature and instincts set out various functional ends connected to my wellbeing, means that these things are *good for me*. But according to Korsgaard it *also* serves as a defeasible reason for me to take those ends to be good *absolutely*. Korsgaard glosses ‘good absolutely’ a couple of different ways: it is ‘good for everyone’ or ‘good (or at least not bad) from every point of view’ or ‘part of a universally shared or common good’ (pp. 134-136). The reason is defeasible insofar as there may be external considerations (involving other agents) that block the bootstrapping move from tethered good to absolute good. So this is not an appeal to a mysterious conception of transcendent or objective value but rather an assimilation of ‘universal’ or ‘absolute’ good to the intersubjective, common good.

Taking the good-for-me as a good absolutely in this way (or, in the more fictionalist language Korsgaard sometimes uses, treating it ‘as if’ it were a good absolutely) involves treating myself as (or ‘as if’ I were) an end-in-itself. Korsgaard speaks of such ‘treating’ or ‘taking as if’ as our way of ‘setting a value on some ordinary end of inclination’ (p. 143). This is a familiar constructivist line of thought: our acts of valuing are what make things absolutely valuable.

Second, if I am rational I also test my maxims against the categorical imperative, and (if they pass the test) regard them as a *law* for others and myself. In doing so, I effectively recognize other rational agents’ ability to make claims on me, too. Thus it is here in the second ‘moment’ that our capacity as *autonomous* beings to recognize the autonomy of others – and legislate in light of it — plays a key role. For Kant himself, the possession of this capacity is what confers moral standing and membership in the moral community.

Now the key move: Korsgaard agrees with Kant that the conscious rational ability involved in the second moment of choice can be the basis of moral standing. However, she thinks that there is also something in the *first* moment that can be such a basis – and that this first moment is something we share with non-human animals. Just like me, my dog has desires which, when satisfied, produce the sorts of happiness and wellbeing of which dogs are capable. In other words, there is a *good-for* my dog — a functional end that is ‘tethered’ to the kind of being that he is. But, just as in my own case, the fact that there is a good for my dog – including the good of life itself — serves as a (defeasible) reason for *me and everyone else* to take that good to be (or to ‘treat it as’) a good absolutely. It is thus a reason for treating my dog as an end-in-itself, a being with moral standing.

Naturally, dogs can’t ‘take’ their goods in this way, or have reasons for doing so, or treat themselves or others as ends-in-themselves – all of that requires higher-level rational faculties. But *we* who can do this are justified in taking *their* tethered goods to be part of the common, universal good. In fact, Korsgaard says, we are not only justified but *required* to do so. But if we *ought*

to treat dogs as ends-in-themselves, then they are our fellow creatures — members of the same ‘moral community’ (p. 148). As a result, we must put in place ‘moral laws whose protections extend to the other animals’ (p. 131).

3. The bootstrapping premise

That is a sketch of a very complicated argument. Before trying to reconstruct it in more detail, I want to emphasize one key point – namely, that for Korsgaard:

[i]t does not follow *directly* from the fact that something is good *for* someone in particular that it is good absolutely, and that anyone has reason to promote it. (p. 139, emphasis in original)

In other words, simply having a tethered good doesn’t make me an end-in-itself. Some rational, self-conscious being also has to *take* the fact that it is a good for me to justify treating it as a good for everyone.

When we do this with respect to our own goods, we take ourselves to be ends-in-themselves — we are ‘valuing ourselves’ as such (p. 137).

It is as if whenever you make a choice you said, ‘I take the things that are important to me to be important, *period*, important absolutely, because I take *myself* to be important.’ So by pursuing what is good for you as if it were good absolutely, you show that you regard yourself as an end in itself, or perhaps to put it in a better way, you make a *claim* to that standing. (p. 139, emphasis in original)

Here it sounds as though self-valuing is fundamental, and that we take goods for us to be good absolutely ‘because’ we value ourselves. But elsewhere Korsgaard makes it clear that these are two sides of the same coin, and what it *is* to value ourselves as ends-in-themselves is to treat our tethered goods as goods absolutely:

The way that we value a creature for her own sake, rather than merely as a means, is by valuing what is good for her, in the final sense of good, *for its own sake*, or *just because it is good for her*. (p. 136, emphasis in original)

Let’s grant that we *can* do this. Why *do* we or *should* we do it? How does the argument go from the ‘can’ to the ‘is’ and the ‘ought’ here? At this point, Korsgaard makes a key bootstrapping move. I take the *bare fact* that there is a good for me – some ends that I want and care about — to justify and oblige myself (and others) to treat that good as (or as if it is) a good absolutely. Moreover, if asked *why* I take the fact that I have a desire for X to justify treating X as a good absolutely, I will have nothing more to say. Such taking is fundamental:

When I make the original choice, when I decide that my desire to grow vegetables is a reason for me to set a *value* on having a garden, I have no other reason for taking my end to be good absolutely, than the fact that it is good *for me*. So I am deciding to treat my ends as good absolutely, simply because I am a creature with a final good. (p. 144, italicized emphasis in original, underlining added emphasis added)

Korsgaard clearly regards the austere commitments involved in this move as an advantage. It avoids appeal to mysterious kinds of inherent value, and thus obviates the need for an account of what that value is and how we recognize it. These were always the Achilles' heels of Regan's theory. As we will see in a moment, however, this refusal to supply a further reason for taking a tethered good as a good absolutely may be a source of a different kind of weakness.

4. The Master Argument: stagewise

I'll now try to reconstruct the master argument in more detail. It comes in two stages: the human stage and the non-human stage. (Note: here I am setting aside the defeasibility of the bootstrapping move just to make things simpler.)

4.1. Human stage

- (1) If we treat a tethered good as good absolutely, then it is good absolutely (boot-strapping premise);
- (2) We treat our tethered goods as good absolutely (empirical premise?);
- (3) So, our tethered goods are good absolutely (1, 2);
- (4) If our tethered goods are good absolutely, then we are ends-in-themselves with legitimate moral claims on all rational creatures (conceptual truth);
- (5) So, we are ends-in-themselves with legitimate moral claims on all rational creatures (3, 4).

(1) is controversial, as we have seen. A non-constructivist will insist that our 'taking' and 'treating as if' do not (or not always) have the power to bootstrap tethered goods into absolute goods. But these debates are familiar and don't bear specifically on the issue of whether non-human animals are ends-in-themselves, so I will set aside discussion of them here.

There are also questions to raise about (2). It is clearly empirically true, given the meaning of 'absolutely' that we are working with here. But *why* do we treat our tethered goods in this way? As we have seen, Korsgaard says that it does not follow directly from the fact that someone has a tethered good that it is a good absolutely. She also says there is 'no other reason' to treat what is good for us as part of the universal good than simply that it is good for us. But at first blush, it is not clear how this counts as a reason at all. If someone asks:

Why do we treat an X as part of Y – what is our reason for doing that?

an informative (not to mention probative) answer will say more than:

Because it is an X.

What the questioner wants is information about *why* being an X (a good for us) makes something worthy of being treated as part of Y (the universal good). But that is the sort of answer Korsgaard thinks we can't supply in these cases. We just do treat an X in this way; there is 'no other reason than that'. This is the bedrock 'is' at the bottom of the constructivist 'ought': another familiar if controversial move which I will return to in a moment.

4.2. Non-human stage

- (6) There is 'no difference' between us and the non-human sentient animals with respect to having tethered goods (independent argument, ch. 2);
- (7) Necessarily, our tethered goods are no more or less important than those of the non-human sentient animals (from the concept of a tethered good, ch. 3-4);
- (8) If we treat a non-human sentient animal's tethered good as good absolutely, then it is good absolutely (1, 7);
- (9) If there is no difference between us and the non-human sentient animals with respect to having tethered goods, and if we treat our tethered goods as goods absolutely, then we ought to treat non-human sentient animals' tethered goods as good absolutely (parity argument);
- (10) So we ought to treat non-human sentient animals' tethered goods as good absolutely and thereby make them good absolutely (2, 6, 8, 9);
- (11) If we ought to treat non-human sentient animals' tethered goods as good absolutely and thereby make them good absolutely, then non-human sentient animals are ends-in-themselves with legitimate moral claims on all rational creatures (from the concept of a final end);
- (12) So, non-human sentient animals are ends-in-themselves with legitimate moral claims on all rational creatures (10, 11).

We have already considered Korsgaard's reasons for thinking that (6) and (7) are true. (8) is an applied version of the bootstrapping premise in (1), and raises at least as many questions.

What I want to focus on however, is the parity argument in (9). Again, Korsgaard says there is 'no other reason' for taking my tethered good to be good absolutely than the fact that *it is a tethered 'good for me'*. If this is right, then 'all we have to do is generalize' in order to get to a principle that covers everyone else and non-human animals too: '[T]hat principle requires that we should take the ends of beings who have a final good to be absolutely valuable' (p. 144). In sum:

[A]ll we have to go on is that some things are certainly good-for or bad-for us. That then is the starting point from which we build up our system of values – we take those things to be good or bad absolutely – and in doing that we are taking *ourselves* to be ends in ourselves. But we are not the only beings for whom things can be good or bad; the *other animals are no different from us in that respect*. So we are committed to regarding all animals as ends in themselves. (p. 145, emphasis added)

Let's suppose that Korsgaard's opponent feels pressed by this parity reasoning. Is there a non-arbitrary difference he could point to that would allow him to resist this move? I can think of two.

First, he could allow that (6) as stated is true: all sentient animals have tethered goods. But he could also plausibly suggest that the *reason* (2) is true – that is, the reason that I take my tethered good and (X) to be part of the good absolutely (Y) — is not just because it is *an* X, but because it is *my* X (that is, it is good for Chignell). In other words, he could (plausibly, I think) say that we each treat our tethered human goods (of existence, health, safety, opportunity, and so on) as part of the common good because, well, they are goods *for us*. Interestingly, as we have seen, Korsgaard herself seems to articulate the premise this way in places: 'I have no other reason for taking my end to be good absolutely, than the fact that it is good *for me*' (p. 144, emphasis in original).

But if that's right, then (9) can be replaced with

(9#) If there is no difference between us and the non-human sentient animals with respect to having tethered goods, and if we treat our tethered goods as goods absolutely *because they are our tethered goods*, then we ought to treat non-human sentient animals' tethered goods as good absolutely (parity argument).

Parity arguments are tricky, but this one now looks invalid: the italicized phrase adds something to the antecedent that makes the conclusion impossible to draw. Real parity would require a consequent like '...then the non-human sentient animals ought to treat their tethered goods as goods absolutely *because they are their tethered goods*'. But, unfortunately, that's not something the other animals can do.

There is a second difference the objector might point to in an effort to resist the parity argument. Recall that there are two sub-moments in Korsgaard's 'first moment' of a decision: the having of a tethered good, and the taking of that fact as a defeasible reason to treat it as absolutely good. Korsgaard admits, as we have seen, that the latter 'does not follow *directly*' from the former (p. 139). But then the logical space between these two sub-moments is a natural place for the opponent to drive a wedge. He can agree that some aspect of this first moment of a decision is a proper basis of moral standing. But he can also insist that it is *not* the mere having of a tethered good but rather the *ability to take* such a tethered good as an

absolute good – to treat it as if it were absolutely valuable — that is the proper basis. So (6) is true, but this is not:

(6#) There is ‘no difference’ between us and the non-human sentient animals with respect to the ability to make our tethered goods into goods absolutely.

Although Korsgaard explicitly does not agree (see sections 8.5.1-8.6.1), my sense is that this second objection and the constructivist picture make good bedfellows. We rational, self-conscious creatures are the ones who have the boot-strapping ability to treat tethered goods as good absolutely, and thereby make them so. *That* capacity, rather than our mere having of tethered goods, seems particularly significant from a moral point of view. It also seems like one natural place to draw a line between those who are fully part of the moral community and those who aren’t. Unfortunately, it is also a line that clearly divides us from the non-human animals.

Given these problems with the parity argument in (9), I find it hard to see how the Master Argument establishes the strong conclusion that:

when we consider the reason why Kant thinks we *must* claim the standing of ‘end-in-itself’ for ourselves, we will see that we *must* claim that standing for the other animals as well. (p. 130, emphasis added)

5. Conclusion

Fellow Creatures is a sophisticated application of a Kantian constructivist framework to the question of how we treat other animals. It offers an original argument for the holy grail among Kantian animal advocates – that is, a way to extend the Formula of Humanity to *all* of sentient animality. Unlike Tom Regan’s Kantian argument for this conclusion, Korsgaard’s admirably avoids invoking a mysterious kind of ‘inherent value’ or an obscure ability to recognize it. It is thus a new and welcome addition to the growing collection of arguments against the ongoing speciesist mistreatment of animals in our society.

However, Korsgaard’s Master Argument does crucially rely on the controversial constructivist idea that *we* make certain outcomes (and selves) absolutely valuable by valuing them in a certain way. It also invokes a parity argument that, I have suggested, even Korsgaard’s fellow constructivists can reasonably reject. Such an objector could suggest that we treat our tethered goods as good absolutely not merely because they are tethered goods, but because they are *our* tethered goods, and thereby also the good for beings who are capable of conferring absolute value on things. If that’s right, then there is a non-arbitrary difference on which to rest a moral distinction between us and other animals. In other words, the objector could insist that the constructivist ability to confer absolute value on tethered goods is a morally significant ability — one that plausibly gives us (and our ends) a special kind of moral standing. But that is just Kant’s own position again, and not the holy grail that Kantian friends of animals were hoping to find here.

Korsgaard has given us a deep and honest book — one that is well worth reading. If my reconstruction of her Master Argument is correct, however, then there is still work to be done by those who seek a Kantian basis for direct moral duties to *all* of sentient animality.*

Princeton University
chignell@princeton.edu
doi:10.1093/mind/fzz076

ANDREW CHIGNELL

* I am grateful to Christine Korsgaard and participants in Marcus Willaschek's doctoral colloquium at Goethe-University Frankfurt for their comments on an earlier version of this review.