

Further Thoughts on *Fellow Creatures*

Peter Godfrey-Smith

University of Sydney

March 2021

These notes accompany "Philosophers and Other Animals," *Aeon* 2021.
<https://aeon.co/essays/why-korsgaards-kantian-argument-about-animals-doesnt-work>. This version is incomplete (v1) – one section to follow.

My *Aeon* essay about animals and moral philosophy was organized around Christine Korsgaard's book *Fellow Creatures* (2018). This book uses a Kantian framework (or Kant plus some Aristotle) to argue that we have an obligation to treat at least many nonhuman animals with far more consideration than is now common. I argued against some central ideas in the book and sketched an alternative view, one that agrees with Korsgaard that we should change a lot of our behaviors, especially in farming and experimentation, reaching conclusions of this kind through a different, less compulsory, path. In that article I wanted to discuss her Kantian approach without using any Kantian jargon at all, and without getting swallowed up by the details of her presentation. Here I'll offer some further commentary about the book, mostly adding thoughts in support of my main argument, and will also say more about the positive view sketched at the end of the *Aeon* article [this part will come later]. These notes are intended as a supplement to the earlier article, but they have ended up being self-contained enough to make sense, probably, on their own.

Before starting I want to express thanks – in a way that doesn't happen in magazine-style articles – to some people who helped me with these issues. I thank especially Andrew Chignell, who was very insightful and generous in correspondence, and also on the less Kantian side, Lori Gruen and Simon Blackburn.

Notes on Korsgaard's Main Argument

Korsgaard gives a Kantian argument that we should treat (many) nonhuman animals as "ends in themselves" an argument that proceeds from reflection on the nature of rational action in general. The starting point is found in our own everyday actions, especially our sense of ourselves as acting for reasons. "Because of the way in which we are conscious of the motives for our actions, we cannot act without endorsing those motives as adequate to justify what we propose to do." (In the passages quoted in this paragraph she is describing Kant's view, but, as I understand her, endorsing it.) The way that we take these motives to be adequate involves a commitment to the idea that what we seek (when we're being rational) is "absolutely good." A commitment of this kind is "built in to the nature of the kind of self-consciousness that grounds rational choice" (all 8.4.1).

The commitment we have to seeing what we seek as absolutely good is a commitment to the idea that what we seek can be recognized as good from the point of view of others. It's not that everyone has to care about the particular things that I care about, but "if my caring about an end gives me a genuine reason for trying to make sure that I achieve it, then everyone else has [as I see it] reason to value my achieving it as well" (8.4.1). Being good in a way recognizable to others, my achieving what I aim at becomes part of a *shared* good. This commitment is implicit in our own choices and our own investment in reasonableness. It is something we, in a sense, already think – "we think that our achieving our ends is good from the point of view of others and not merely good-for-us" (8.4.3). So we not only think that we have good reason to pursue the projects we do, but also "expect others not to interfere with that pursuit without some important reason for doing so, and even to help us pursue them should the need arise" (8.4.3).

This is presented as a tacit recognition, within our own attitudes, of the status that rationally sought goals have. From there, we are supposed to recognize that we should have the same attitude to the projects and choices of others. This doesn't mean that we have to be as invested in others' projects as we are in our own, but it entails a basic respect for others' projects and goals.

Korsgaard presents this as a central thread in Kant. *Fellow Creatures* is novel in applying this kind of argument to our relationships with animals as well as other people. Animals can't be part of a community featuring a reciprocal respect for autonomy, as

people can, but Korsgaard thinks the considerations above can be used to constrain our behavior towards them, too. We should recognize that other animals pursue goods that we should recognize as real, and hence we should not interfere with their activities and projects without good reason.

The main objection I have to this part of the book is directed at some of those basic moves derived from Kant, rather than the application of these ideas to animals. I think there is a structural problem affecting the idea of "shared" and "absolute" good as employed in the book.

Here is a presentation close to the version in the *Aeon* article. Suppose I decide to do something. I think I have good reason. Because I think this, I think that others will, or could, also see that it makes sense for me to do it. They will see that if they were in my shoes, they would want to do something similar – more accurately, not if they were in *my* shoes, as it shouldn't matter which particular person is involved, but in shoes *like mine*. In a way, the goodness of what I am after, along with the reasonableness of what I am up to, should be visible to everyone. So far, though, there is no reason why I should expect them to put any value of their own on what I am doing, as they may be in a different situation. They will endorse it, I think, for anyone in shoes like mine, but there's no reason yet for them to endorse it beyond that.

Part of the problem comes from the word "absolute" as it is used here. Korsgaard intends that term in a non-metaphysical sense. Something is absolutely good when it can be recognized as good by everyone. Value is constitutively tied to valuers and valuing, so "absoluteness" can't be more than a matter of agreement on all sides – a kind of perspective-independence. But there are two ways something can be recognized as good by everyone. It might be recognized, by everyone, as good for anyone who is in shoes like mine. That does not mean it is recognized as good in a further sense where it becomes part of a *shared* good, a good that everyone has reason to pursue.

I'll now say more about this, and make a connection to the Kantian idea of an "end in itself," a concept that Korsgaard does use when presenting her argument. Korsgaard says that we, through our ordinary choices, claim the status of "end in itself" for ourselves, and then we come to recognize that others – other people and also animals – are also ends

in themselves. One thing I want to do in these notes is say why it doesn't make much difference to express the argument in those terms.

I'm not going to address what Kant himself meant by "end in itself". My topic is the role played by a version of this idea in Korsgaard's argument. What this concept gets us – where it leads and what it motivates – depends on what goes into it.

Korsgaard says that we do, and have to, value ourselves in this particular way – as ends in ourselves. We do this by *taking* what is good for us to be good absolutely, by taking what is good for us as an end of action. When we say that something is good absolutely, what we mean is that it is good from everyone's point of view.

When we say something is good absolutely, what we mean is not that it has a free-floating goodness, but that it is good-for everyone for whom things can be good, in the final sense of good, or good from everyone's point of view. To put this in a somewhat different way now, which will be helpful for the purposes of this chapter, we mean that it can be included in a universally shared or common good, one that we can all pursue together. (8.3.1)

To see other another creature as an end in itself is to see that we should "accord the creature the kind of value that... she necessarily accords to herself" and "therefore see her final good as something worth pursuing" (8.3.3).

The problem I described earlier is seen at a particular place in the quoted passage above. That place is where we go from "good from everyone's point of view" to "shared or common good." The problem arises because the sense in which we, as rational agents, think that our ends will be recognizable as good from others' points of view is a sense that has the context-sensitivity described earlier. I think that others will, or could, see that if they were in shoes like mine, they'd recognize that what I am seeking is indeed worth seeking, also. But what I think they'll see, again, is the way these goals would make sense for someone in shoes like mine. They can see this without regarding what I seek as "included in a universally shared or common good." They *might* see my goals in that way, but nothing about the nature of reasonableness or agency motivates or compels them to. If an end in itself is something whose good can be recognized as good from everyone's point of view – if that is how we reach the concept of an end in itself – then there is no bridge here to the idea that what is good for me, or you, becomes part of a shared good.

Again, I am not saying that this is Kant's own concept of an end in itself, just that this is the concept that Korsgaard uses in her argument. If the *basis for christening* something as an end in itself is that their good can be recognized as good from others' point of view, then this is not enough to make their good into a shared good, or make their interests something we should take into account. Those prosocial attitudes might be reasonable, or reasonable in some cases, but the materials used to establish the "end in itself" status do not deliver it. This is, again, because of that ambiguity in talk about "absolute" good, and things being recognized as good from others' points of view.¹

The following single sentence contains, in a sense, the main claim together with its problem: "The absolute goodness of our ends is a presupposition of rational action." (8.4.2). The thing that might be a presupposition of rational action is that what we are doing makes sense in a way that others might see. What others might, or should, see is why an action suits circumstances like mine. It's not a presupposition of rational action that our ends are, or even might be, part of a shared good.²

Putting it yet another way (I realize I am doing this over and over, from only slightly different angles): if the "end in itself" idea is motivated by noting some tacit commitments of ordinary rational action, then there is indeed a concept of a *coherent goal-seeking agent* we can get to via this road – the idea of someone whose pursuits make sense, from many vantage points, given their situation. But I can recognize you as a coherent goal-seeking agent while also seeing that I have no reason to help you pursue the goals that make sense to you, given those details of your situation. Again, when I talk of

¹ Here is another expression of the idea, in 8.8.3: "I have characterized an absolute good as one that is good from every point of view, and also as one that can be included in a shared or common good which we can all pursue together." Once we note the context-sensitivity of reasons for action, those things she links by "and also as" become quite different.

² Korsgaard at one point makes a comparison with reasons for belief, as opposed to reasons for action. She says, considering a situation where we are looking at evidence bearing on whether a particular person committed a crime, "Reasons for belief must be the same for everyone" (7.3.2). The situation is similar. We might have a situation where I can see that I was in shoes like yours and saw the evidence you saw, I'd believe what you do. But my evidence might be different, and then I might have different beliefs. I might have no inclination at all to believe what you do, even though I can see that *if* I had your evidence, then I probably would. Your conclusion is defensible, coherent, justified in one important sense, but it's not something I have any inclination to take on board.

your "situation" here, I mean not your particular identity, but the qualitative, in principle repeatable, details of your situation.³

The end in itself concept is usually taken to be a good deal heavier than this. It is *supposed* to contain more and do more. It is something like "a being whose good matters absolutely" (a phrase from Korsgaard's "Valuing Our Humanity," not *FC*). But the way that "absolute" goodness is motivated in *FC* prevents end-in-itself status from taking us very far down the road.

I'll discuss one more aspect of this issue (this is now at least a small gear change). Looking back at that passage I discussed in detail earlier, as it goes on from the place I ended the quote, it introduces another element:

When we say something is good absolutely, what we mean is not that it has a free-floating goodness, but that it is good-for everyone for whom things can be good, in the final sense of good, or good from everyone's point of view. To put this in a somewhat different way now, which will be helpful for the purposes of this chapter, we mean that it can be included in a universally shared or common good, one that we can all pursue together. Among other things, putting it that way will enable us to include things that are good for someone, as long as they are not bad for anyone, among the absolute goods. (8.3.1)

Where does the qualifying clause "as long as they are not bad for anyone" come from? Including something like this seems inevitable, as some possible claims about a transition from our taking something to be good to it being absolutely good are clearly too simple. It might be good for me to take some of your poorly guarded possessions, but unless my need is greater than yours, that can't be part of a common good. Here is another passage that is related:

³ My aim here is to discuss Korsgaard's argument in *FC* rather than the entire Kantian picture, but I do wonder whether other parts of that picture are affected by these issues. One is the idea of universalizability. We are supposed to act in a way that would make sense if its principle was followed universally – if everyone followed the same rule. How many of the idiosyncrasies of circumstance might be carried along as part of a universalized rule? (Suppose everyone followed the rule that *if* they are in X,Y,Z.. then they do A.)

But your right to confer absolute value on *your* ends and actions is limited by everyone else's (as Kant thinks of it, every other rational being's) right to confer absolute value on *her* ends and actions in exactly the same way. So in order to count as a genuinely rational choice, the principle on which you act must be acceptable from anyone's (any rational being's) point of view – it must be consistent with the standing of others as ends in themselves. (8.4.4, p. 140)

There are two ways of thinking about this. We can see it, again, as a sort of qualification of the initial move that is supposed to take us from the reasonableness of our own choices to a claim about what is absolutely good – it would not make sense to see myself as entitled by this reasoning to trample over others, so we circumscribe the bounds of the claim in an additional move. Alternatively, the point can be seen as more integral to the argument: I think that others must recognize my reasons as sufficient to justify my choices, but the importance I accord to rationality here, in my view of how others are constrained, compels me to recognize that rationality in others is sufficient to "confer absolute value" on what *they* are seeking, so I must also recognize the value of their projects and operate within that recognition. It's not a matter of qualifying a *prima facie* claim that is leading somewhere unacceptable; the constraint is more intrinsic to the argument. I guess that is more likely to be what's meant, but if so, the problem discussed above returns again. As I abscond with your possessions, I can recognize that the principle on which I act is, in a way, "acceptable from anyone's (any rational being's) point of view" – I know it will make sense when seen from the outside, and recognize that you, with similarly good reason, also have your eye on some loosely guarded possessions of mine. Both sides are rational, both know the other can see good sense in the choice being made, and neither choice is prosocial. If we want to motivate prosociality, we need different resources.

I see *FC* as an admirable book. It is accessible but also philosophically ambitious. Grappling with it has been a rewarding experience. Here are some last comments about it (for now) made standing a little further back. One of the things being attempted in a neo-Kantian project like this is a bridging of self-other divides in the area of choice. Different forays of this kind are seen in different parts of Korsgaard's work – I don't know all that work, but some of it. In *Sources of Normativity*, Korsgaard more readily confronts the problem of rational self-interest, but argues from the public nature of reasons, via

Wittgenstein, that there is a kind of collapse of that challenge. The approach in *FC* is simpler and more minimal. It attempts a bridging based just on an understanding of what it is to make and recognize reasonable choices, and I have argued that this is not enough to take us where Korsgaard wants to go. That leaves open the possibility that an argument of this kind could take us further if other elements are added. One is the philosophical rejection of certain kinds of self-interest I mentioned above. Another (and these can be combined) is to say that the particular thing we are placing value on in our ordinary choices is a feature of ourselves that we must recognize as shared across others, so that valuing our own instance (as we inevitably do) commits us to valuing it more broadly.⁴ There is a kind of compulsory export of concern, based on what we are valuing when we make ostensibly self-interested ordinary choices. The problem with these further moves is the problem of showing that our ordinary actions and choices have anything like this built into them. I doubt that they do. Trying to show they do tends to require a lot of the more elaborate Kantian apparatus that Korsgaard eschews in *FC*.

2. Further Notes on the Positive Proposal

[to come]

⁴ "Interacting with Animals: A Kantian Account" (2011). "The stronger way to make the argument is just to say that because the original act of self-respect involves a decision to treat what is naturally good or bad for you as something good or bad objectively and normatively, the self on whom value is conferred is the self for whom things can be naturally good or bad. And the self for whom things can be naturally good or bad is your animal self: that is the morally significant thing we have in common with the other animals. It is on ourselves *as possessors of a natural good*, that is, on our animal selves, that we confer value. Since our legislation is universal, and confers value on animal nature, it follows that we will that all animals are to be treated as ends in themselves." She also considers here a "weaker" option based on the special status of some natural goods such as avoiding pain.
<https://www.people.fas.harvard.edu/~korsgaard/CMK.Interacting.Animals.pdf>.