

SAPERE AUDE: KORSGAARD INTERVIEW

SA: Professor Korsgaard, thank you for talking to us. First, would you mind explaining the general project of your paper?

CK: The paper is part of a larger project. The way I frame the question in the larger project is as a question about how there come to be any such things as values. I ask the question from a naturalistic perspective, so I suppose that the world is roughly the world that science describes to us. What is already in the world are the sort of things physics and biology says are already in the world. It doesn't already contain, among its fundamental parts, values. Because I believe values depend on valuing – rather than the reverse. You can understand the relation between values and valuing in either of two ways: you can think that when we are valuing there are values out there and what we are doing is responding to them, or you can think that our valuing somehow produces values. I think the second thing. The question is: how does this happen? The paper is the part of a larger project aimed at showing why there is such a thing as a final good.

SA: One thing you talk about in your paper is the idea of your body welcoming or rejecting things that come to it. Can you clarify what you mean by welcoming and rejecting?

CK: It is a little hard to do that because I take those to be very fundamental attitudes built into the way in which we experience the world. They are built into the nature of consciousness itself. Some people have a picture of consciousness as just a source of information about the world. You look around the world and you are aware of it. You gather up all this information, and then having all that information in hand, you have to decide what to do. Now if you think about it, you can see that that cannot possibly be how it works with the lower animals, in whom consciousness first evolved, because they didn't have the equipment for figuring out what to do. So I think instead that what we in fact have is a fundamentally evaluative form of consciousness, meaning that things strike us (and all the animals) as being attractive or aversive in various ways. Furthermore, things strike the animals as being attractive or aversive in quite particular ways that guide the animal as to what to do. In a sense, I think, when a lower animal sees the world, he or she sees the world in a certain way with instructions already packed into it. For example, a predator is a thing to flee from, whereas an offspring is a thing to care for. With foods of various kinds, something smells a certain way, and that

makes it good to eat. The perception comes with instructions in it. What is different with the human being is not that we have lost that form of consciousness altogether, but that we are able to step back from it and assess the instructions that we're being given and decide whether they are good ones to act on. I take the attractive and repulsive parts of experience to be a fundamental part of the way we relate to the world.

SA: You have an Aristotelian version of pleasure and pain in which they are responses to experience rather than sensations, and that seems right to me. It seems like they might be somehow analogous to thick concepts in McDowell's and Williams' sense, in that we cannot really understand them without understanding their normative character. It seems to me that we can also have pleasure and pain sensations without the normative element in some cases. Can you talk about these cases in light of what you are doing in this paper?

CK: Well, I am not sure it's exactly right to say there can be pleasures and pains without the normative element. But here is a case where the normative element is, as it were, silenced. I said several times last night that pains and pleasure were not the objects of experience. And I think that is true in the first instance. Of course, when you are in this state of pain or pleasure you can back off and think about it, and then it becomes a kind of object of experience. And one thing that human beings can do, that I think the other animals can't, is in a distanced way enjoy and appreciate having pains and pleasures. For instance, think about enjoying a scary movie. The thing that is going on on-screen is appropriate to react to in a certain way; and there, sitting there in the dark movie theatre, you are cringing, or tensing up, or feeling terrified, and reacting in just the ways you think the hero should be reacting. But at the same time, the normativity is cut off, so that you are just experiencing it and enjoying the experience. So in that way, I think the normativity can be cut off.

SA: Last night you said that you cannot treat individuals as they are at just a single time. You have to treat them as they are overall. How does this view apply after someone's death?

CK: I think the attitudes we owe to people cease to be attitudes that we owe to anything physically identifiable, at the stage you are talking about, but we do continue to owe certain attitudes towards the person considered as a whole. For instance, sometimes I am made morally nervous by highly fictionalized accounts of historical characters. I think, "Wait a minute! You don't just get to play around with that person any old way

you want!" There is someone who was there and in a sense is still there, and we owe certain attitudes to them; the fiction might be disrespectful or even count as lying about the person.

SA: How would you go about regarding such a seemingly timeless presence cognitively? You have to think about things one at a time, you can't just have everything going on at once.

CK: I am not sure I would see a case in which this would apply.

SA: I'm thinking about the overall make-up of how you think about something. When a person is born, in order to treat that as timeless, you would have to say you know everything about them, including their future, and that's untenable.

CK: I don't think you have to assume you know everything about them, because, at least for our purposes, a lot of things about them aren't settled yet. I suppose this might raise some worries about determinism, according to which everything about people is settled, but we do stand in time, and from our perspective many things about people are not settled yet.

SA: Your judgment about something being good or bad is based on the identity of that thing. Animals in virtue of being animals don't choose that identity. They have certain relational goods or bads for them, and humans have goods and bads for them as humans. And then we can add all these identities on top of that, for example the assassin. How do you want to frame goods and bads just in virtue of being human? Is it Aristotelian, how do you judge something being good or bad, just from a human standpoint?

CK: The good for human beings is complicated, because of the ways in which we construct our own identities, which the other animals don't. There are various roles, relationships, occupations that we take on and these are normative for us. You decide to be someone's friend, or you decide to have a child, or you decide to take a certain position, and these things all come with things you have to do, in order to have that as part of your identity. And then what's good for you are the things that enable you to fulfill those roles and relationships well. So you do things that make it the case that certain things are good for you. It is also true that also just insofar as you are a human being some things are good for you, including what enables you to do what I just said: take on roles and relationships and so on. For instance, when we think about what is good for people in a political context, we think about making a world in which people

are free to choose their own identities and can enact those identities without clashing with each other and themselves.

SA: The way I perceive what you are trying to say is that the evaluative good collapses into the final good. Each of those things you have mentioned, i.e. being a friend or a parent, is evaluative.

CK: Yes, that is my whole story: your final good is—in a way—just your evaluative good, insofar as being conscious puts you in charge of it. There is a way in which the difference between a plant and an animal is that the plant just grows, while the animal is in a certain way in charge of his own survival and reproduction. He has to do things to ensure it, and the way he responds to the world has evolved to enable him and to motivate him to do things that ensure it as far as possible. Being in an evaluatively good state becomes the animal's goal, and that's what makes it the final good. The general thesis is that it is because there are such beings there is a final good, because there are beings who are in charge of their own evaluative goodness. What makes it the case that there are objective facts about the good is that there are objective facts about when an animal or a person is or is not well functioning. Although in the case of a person, many of those facts are rooted in the person's own choices, and therefore, in how he sees things. If you take on a certain practical identity, say you are friends with someone and you see friendship a certain way, and hopefully your friend does too, then there are facts about doing well as a friend that are dependent on how you understand it, and to that extent, there is a subjective element to it. But there are also objective elements about whether the way in which you construct yourself is meeting your human needs. And—of course—for the animal there are objective facts about the good because there are objective facts about whether the way an animal acts is conducive to its survival and reproduction.

SA: You said that we are faced with the task of constructing a state of affairs that is good for us all. I am wondering if that involves creating an environment where people are free to do those evaluative things, or are you thinking of something else?

CK: I have two things in mind: first, I thought, after the discussion yesterday, that people were thinking that I wanted to get more moral content out of the shared goods idea than I do; it is just one idea among many others that you can use to get moral content. The two things I had in mind are these: first of all there is what you just said: there is a way in which we can construct the political environment with a view to

having it be good for everyone, with a view to realizing the good for as many people as possible. The second concerns a more technical issue in moral philosophy. Utilitarians think of the good as aggregative; and there are two problems with that from my point of view. The first arises from the fact that I think goods are relational: it doesn't make any sense to add relational goods because the result is not good for anyone. My good plus your good—whose good is that? So I don't think aggregation makes conceptual sense. The other problem is that if you can aggregate, not only can you add, but you can subtract, and that is why you get the problems utilitarianism gives rise to about examples where—for instance—if everyone enjoyed watching you getting eaten by the lions, the overall good might include you getting eaten by the lions. Everybody realizes these are bad consequences, and utilitarians work very hard to show that this wouldn't happen, but it remains the fact that such results are in principle possible. So there are reasons to worry about aggregating goods, though there are many intuitions that seem to favor the idea, intuitions according to which we can do the most good by doing certain things, even when we are not clearly doing good for anybody in particular. For instance, many people believe that given the choice you should save five lives rather than only saving one. One explanation of why you should do that is because it does more good. That's not the only explanation, but it is a tempting one. Here is one of a completely different kind: I have one effective dose of painkiller, one person has a mild headache, another is in agony, to whom should I give it? It looks as if I should give it to the person in agony, and one explanation is that that does more good. Somebody who doesn't believe in aggregation has to find a way of thinking about those cases that's different. I think some of those cases can be handled by the idea of looking for shareable goods, not all of them but some of them, and I also think the idea of shareable goods makes sense of some cases where our intuitions about aggregation are fuzzy. Suppose you say, well, I could do something that is going to hurt just one person, but its going to help all these other people. If you're thinking of an example where this is going on in a town, in a community, you might have the intuition that this is a way of doing the most good. On the other hand, suppose the case is one where you can push a button and this will hurt somebody in Africa, but it will help many other people across the globe who have nothing to do with one another. Can this be explained by the idea that this way you are doing the most good? When you have the idea of shared goods in hand, you might think there is something that is good for the community, which would

explain why you would do something that is better for more people in the community, but wouldn't give you the conclusion that you should do something that is better for some random group of people scattered across the globe, because there is not a unit for whom that is good.

SA: I have a question that relates to *The Sources of Normativity*, specifically the fourth lecture, where you talk about the publicity of reasons. You talk about how somebody says something to you that is offering you a reason and you can't hear it as mere noise, you hear it as a reason; you can accept it or reject it, but you can't go on as if you hadn't heard it. It seems to me like you can have cases where your practical identities make it that someone is trying to offer you a reason, but you can't hear it as a reason. An example I am thinking of in this context is the Stanford prison experiment: you have the guards and you have the prisoners, it seems like at least one way of analyzing that situation is not the prisons are obligating the guards, and the guards are rejecting it, but rather, the guards are simply hearing the prisoners pleas as mere noise – as not reasons at all; and so they can go on in the same way. What I am wondering is how on your account you would analyze or interpret this case differently.

CK: Certainly we have ways of shutting out the force of the reasons we are given, and that would be what I would be inclined to say here. It is not really that it is mere noise, but—one way we shut things out is that even though we don't actually hear it as mere noise, we tell ourselves that is what it is: "There they are whining again," we say; trying to reduce their complaints to mere noise.

SA: The way I was trying to construct the case was: they are certainly not hearing it as mere noise, but it seems like something about the situation—you are making a kind of Wittgensteinian argument, so you are going to say: giving reasons only makes sense in the right sort of context. In my example, what I am suggesting is that, for the guards, prisoners are not the sort of people who can be reason-givers. That is my fundamental worry, it seems like it's a context where it has been established were these people aren't allowed to give reasons. Them giving reasons is just not part of what happens in this situation, as hard as the prisoners might try.

CK: I don't know the experiment so I don't know what is reported as going on with the guards, so can you fill that in a little bit?

SA: It was an experiment where a psychologist at Stanford took a group of students and separated them into two groups: one group the guards, the other the prisoners. The

experiment had to be shut down early because the guards got so vicious and demeaning towards the prisoners, even though the group was all peers; but once you establish that separation and gave the guards carte blanche, short of something illegal, then you have this problem where the prisoners are pleading for mercy, and the guards are doing all sorts of horrible things to them.

CK: Its complicated isn't it? You can't be vicious and demeaning to just anything, you can only be vicious and demeaning to something that is conscious, sentient, or rational. So there is a way in *those* forms of cruelty are complicated precisely because they depend, in a weird way, on thinking of the other person as a fellow rational being or a fellow sentient being, or whatever it might be. And that makes me think it is not *merely* not hearing what they say as reasons. Because if they were only hearing it as the squeak of the machinery, that would not be an occasion for being vicious and demeaning.

SA: So you have something where there is something behind the practical identities. There is this shared human identity, even where you have this context, where these people don't get to give me reasons, there is still something behind that that says, well yes, they actually do. And so you have to shut *that* out.

CK: Yes.

SA: One of the worries you have with utilitarianism is that they have these private mental states that are good are bad in themselves and you can have them, one of the problems is that they are reifying the locus of the good into these mental states. And then you have objective conditions can be good or bad for us, my first thought here was that you are doing the same sort of thing as the utilitarian, and you are reifying the good into objective conditions, but on reflection I don't think that is right. Can you talk about why?

CK: When I say that the good is relational, I mean that there is an objective condition that is good, but good in virtue of the way it is related to someone. So it is not that the relation itself is reified and becomes the good, it is the object of the relation that becomes the good, but it is not the good by itself independently of the relation that bears to someone. It is good in virtue of the relation it bears to someone.

SA: You have a bigger project that asks where values come from, and then there is actually applying them to specific cases, can you talk about what a moral decision in general is going to look like on your view.

CK: All of my ideas are in a Kantian framework. I haven't written that much about specific practical questions because I basically think if there is anything like a technique for answering moral questions, it is doing the kind of thinking that's suggested by Kant's idea of treating humanity as an end in itself. There are two practical questions I have written about. There are a couple papers in which I have written about lies, and why you shouldn't tell them, and how respect for persons requires you not to tell lies. And I have also written in recent years about animal rights, and how that can fit into a Kantian account. That is related to the paper you heard yesterday, because the moment when, as I described it, you sympathize with yourself, that transforms the view you take of the world insofar as you are a conscious animal, into a world of reasons, because at that moment you decide to treat the motives you naturally have as reasons. I think at that moment, you are also—in effect—rendering the good of animals normative. Because animals are also among the beings who, in the same sense we do, have a final good.

SA: What was John Rawls like as a teacher?

CK: First of all as a teacher, there was nothing fancy about him, he wrote out his lectures and he read them. They were absolutely wonderful lectures, even without anything fancy going on. He talked about a lot of figures in the history of philosophy and he always told us that when you study a philosopher, you should make the best possible case for his or her view because nothing less than the best possible case for his or her view could possibly be worth objecting to. If you object to a cheap version of the view, then you haven't done anything; you have to go on to come to terms with the better version of the view anyways. The idea is to learn as much as possible from the figures you study. It was very inspiring, and many of his students went on to study the history of philosophy, because they loved that way of approaching it. You could see in his own work how much he had learned from studying the history of philosophy, and how that enabled him to be who he was, even though his main work was not historical. If you read Rawls' early papers they are loaded with all these historical footnotes; in those papers, it's on the page, what he is learning from whom. Later that disappears, so you have to know more things to spot how the history of philosophy is inspiring the ideas in *A Theory of Justice*, but it is there. As an advisor, I think everyone noticed that he treated you very much as a colleague, he has footnotes to his students all over the place in his work, you were very much his equal, he gave you a lot of independence—

sometimes a little bit too much. He wasn't a great detail advisor, in the sense that he didn't sit and go over your work line by line. However, there are certain moments, when you are at an intellectual crossroads and you could go either this way or that way, and he would always spot that and say, "Go that way." At those kinds of moments, he always got that right. I should also say that this was at the point at which things were changing for women in the profession of philosophy - they haven't changed as much as we had hoped, but they were changing. Rawls was an advisor to an enormous number of women and I think this was partly because he was so egalitarian. Feminism was big in those days, so a lot of professors bent over backwards to be encouraging to woman: Jack didn't bend over backwards for women because he didn't need to. He already regarded us as equals. Women were very comfortable working with him. You didn't feel like someone was trying to help you, you felt like someone accepted you.

SA: Why don't you think more women aren't in philosophy?

CK: This is a question the profession has been struggling with the whole time I've been in it. One thing people used to always mention was the lack of examples, so we thought that once there were more women professors, there would be more women entering the profession. So far, that has not been borne out: Harvard has a very good record of tenuring woman professors, for a while three out of twelve professors were women, and still we are not getting more women undergraduate majors than we used to. And that surprised everyone; we really didn't see why that was. Another standard story is that there is a sort of debate atmosphere in philosophy—people argue and they interrupt. People think that women are not trained to interact with people like that, and therefore we are at a disadvantage. Women think it is rude to interrupt, some of them think it is rude to argue, although that is the wrong sense of "argue," so there is that kind of explanation - more generally, that there is a certain aggression required for getting on in philosophy. Although the women who did make it into the profession reduced that aggressive atmosphere considerably from what it was like when I first entered the profession. Another common explanation is the claim that women and minorities don't go into philosophy because philosophy is somehow not speaking to their concerns. So a lot of efforts have been made at developing kinds of philosophy that are supposed to speak to the concerns of women, or people of color. I find that idea depressing or wrongheaded because philosophy is supposed to be speaking with the universal voice; it is supposed to be about the human experience as such. It shouldn't

be the case that the way you get to have a universal audience, is to address particular concerns. I, at this point, don't feel like I know the answer to this question. It may be that the one I find most tempting is the one about philosophy's aggressive style.