REVIEW OF *THE SOURCES OF NORMATIVITY*

Korsgaard wants to provide a foundation for morality that is binding for every rational being, but without resorting to the claim that moral facts just exist in the world in some mysterious way.

The most interesting thing about her theory is how she grounds morality in **identity**. If an important part of my self-conception is (for example) that I'm a good parent or friend or citizen, I can't mistreat my child or friend or fellow citizen without challenging my own identity: mistreating them would force me to acknowledge that I'm no longer the kind of person I saw myself as. We're often very motivated to avoid such challenges to our self-conception.

But Korsgaard is searching for a universal moral law, so she needs to ground it in an identity that everyone shares. This identity, she proposes, is "Citizen of the Kingdom of Ends"; that is, a rational being whose goals are valuable. I find her argument somewhat hard to follow due to its meandering presentation, but perhaps the clearest overview is near the end:

> Guided by reflection, we may be led to see that our tendency to treat our contingent practical identities as the sources of reasons implies that we set a value on our own humanity and so on humanity in general. This realization leads us to the moral principle of valuing humanity as an end in itself.[1]

Instead of offering us a way to say *X is objectively wrong*, Korsgaard's theory is meant to offer us a line of argument we can use to *convince any rational person to believe that X is wrong*. She believes "any reflective agent can be led to acknowledge that she has moral obligations"[2]. As I understand it, the strategy relies on drawing out some latent contradiction between believing that your own goals are worth pursuing and believing that you have no moral obligations.

I think there are a lot of weak points in the book, but I'll focus on two here.

[1] Christine M. Korsgaard, *The Sources of Normativity* (Cambridge ; New York: Cambridge University Press, 1996), 250.

[2] Ibid., 125.

1. Ambivalence about nihilism

Technically, Korsgaard doesn't claim that her view proves you ought to care about everyone's interests. She only claims to show that you *either* ought to care about everyone's interests, *or* adopt total "normative skepticism" and not care about even your own interests. My own preferred argument for morality involves a similar dichotomy: I think the arguments for reductionism about personal identity show that *either* you have a reason to care about everyone's future pleasure and pain, *or* you don't even have a reason to care about your own future pleasure and pain. On my view, though, these are not equally valid options; the absurdity of the nihilistic option is meant to help you see that the moral option is objectively correct.

By contrast, given Korsgaard's metaethical position, it seems like she would have to say the choice between morality and nihilism is entirely arbitrary. I see this as a weakness because I think we can directly perceive that there is something *non*-arbitrary in the choice. Although it's plausible that for many of the things we value, such as art and music, their value derives from our attitude/reactions to them, I do not believe this is true of (for example) the subjective experience of pleasure or joy. If I imagine a person who experiences those feelings in the exact same way I do, yet does not believe or act as if those feelings have value—perhaps makes every effort to avoid experiencing them—it seems to me that the person must be in some way mistaken or confused. That person's failure to value joy does not render their joy valueless. But on Korsgaard's view we would have to conclude that it does.

2. Failure to establish universality

Let's assume this part of the above quotation is a sound argument (though I don't think it is):

> our tendency to treat our contingent practical identities as the sources of reasons implies that we set a value on our own humanity

Then Korsgaard needs to establish that the next bit follows from that:

> and so [we set a value] on humanity in general

But her argument for this seems very weak. She relies on the claim that reasoning is an inherently public or social activity, and that simply learning about someone else's reasons automatically leads us to acknowledge those reasons as having some normative force:

> If I say to you 'Picture a yellow spot!' you will. What exactly is happening? Are you simply cooperating with me? No, because at least without a certain active resistance, you will not be able to help it. Is it a causal connection then? No, or at least not merely that, for if you picture a *pink* spot you will be mistaken, wrong. Causal connections cannot be wrong. What kind of necessity is this, both normative and compulsive? It is *obligation*.[3]

That's an, um, *creative* way to use the word "obligation". But the following is true to some extent:

> All I have to do is talk to you in the words of a language you know, and in that way I can force you to think.[4]

And thus other people can force us to consider *their* reasons in our own decision-making. Korsgaard thinks that when this happens, it's natural and normal for us to consider those reasons as mattering by default:

> We do not seem to need a reason to take the reasons of others into account. We seem to need a reason not to.[5]

Okay. But what if we say that a sufficient "reason not to" is simply: it doesn't affect *me*? Korsgaard thinks the other person can then ask us to imagine ourselves in their position, and we will automatically (perhaps involuntarily) do so, and will thus feel the force of their reasons ourselves.[6] And she seems to think we'd be involving ourselves in a contradiction or inconsistency if we believed something like *if I were you, X would be a valid reason* at the same time we believed *X is not a valid reason*.

But that's not necessarily true. Reasons can be agent-relative. If we're playing chess, for example, I acknowledge that you should try to checkmate me, but I also believe I should try

[3] Ibid., 139.
[4] Ibid.
[5] Ibid., 140–41.
[6] Ibid., 142–44.

to stop you from checkmating me. There's nothing inconsistent about that. Are moral questions like that? Can it simultaneously be true that you should try to benefit at my expense, and that I should try to benefit at your expense? Most of us would answer no, but some people say yes, and Korsgaard's arguments seem insufficient for showing that the latter group is making any sort of error.