Response to Gary Watson and Hilary Bok

APA Author Meets Critics on Fellow Creatures

January 10, 2020

I'd like to thank Gary Watson and Hilary Bok for their generous and helpful comments. Since time is limited, I will set about responding to the specific questions they have raised without further ado. We can take up other questions in discussion.

1. In the book I argue that there are two senses of "end in itself." The sense in which animals are ends in themselves is that, like us, they have a good—things can be good or bad for them—and I argue that because of that we owe them beneficence. The sense in which rational beings are end in themselves is that they are autonomous beings with whom we stand in relations of reciprocal lawmaking—we make the moral law together—and I argue that we therefore owe them respect for autonomy. Since the other rational beings are also animals, and so owed beneficence, this raises a question about how these obligations are related to each other, and a worry that these two sorts of duties might come into conflict, which I discuss very briefly in the book. We might just live with the possibility of conflict between two kinds of duties, but that's not a very Kantian solution, since Kantian ethics is decidedly anti-paternalistic. We might think that autonomy is part of the human good—one of the things that is good for people—but that leaves the possibility of conflict in place. I suggested

instead that our duty to promote the human good takes the form of respecting their autonomy. Gary is worried that respect for autonomy will then collapse into a kind of beneficence.

There is a general question about Kantian ethical theory in the background here, which is how autonomy in Kant's sense—being a participant in moral lawmaking—is related to autonomy in the familiar sense of having control over your own life and actions. They certainly are related, for in Kant's theory there are strong arguments against the use of force, coercion, and deception which have nothing to do with the ends for which they are employed, and therefore there are duties to avoid these methods even when it is supposedly for someone else's good.[1] Now perhaps saying that promoting the good of humans takes the form of respecting their autonomy was a misleading way to describe my third solution. I did not mean to suggest, or anyway I shouldn't have suggested, that the Kantian duties to avoid coercion and deception should be understood as forms of beneficence, or that all duties should be understood in terms of beneficence. I think on the third solution we can still talk about beneficence in its role as just one of our duties.

[1] "The Right to Lie: Kant on Dealing with Evil"

There is a question about whether the duty of beneficence is a duty to promote the ends, the happiness, or the good, of others. Kant describes it both as a duty to promote the happiness of others and a duty to promote the ends of others, and reconciles those (sort of) by claiming that the duty to promote the happiness of others is a duty to promote their happiness *as they see it*. I believe he thinks these things because he also thinks that non-moral good, the good of individuals, is happiness, so he thinks all of these things come together.[2]

I think Kant is wrong about some of this, and my view that beneficence to humans takes the form of respect for autonomy partly depends on the ways I think he's wrong. One of the things he is wrong about is that the non-moral good is happiness, especially if we take happiness to mean something about the satisfaction of your inclinations. (Nietzsche famously said "man does not pursue happiness; only the Englishman does that," and I think Kant got this view about the good from the British moralists whom he admired.) To explain my view I need to sketch my account of what it is for something to be good for someone in the sense of a final good. There is a functional

[2] Kant does acknowledge that we might disagree with someone about what will make them happy, and to cover this case he appeals to another view he holds on the subject: the idea that the duty of beneficence is an imperfect duty. In the *Metaphysics of Morals*, he says:

> "It is for [others] them to decide what they count as belonging to their happiness; but it is open to me to refuse them many things that they think will make them happy but that I do not, as long as they have no right to demand them from me as what is theirs."

So Kant allows for a sort of negative paternalism. He thinks you may refuse to help people to promote ends that you think they shouldn't have chosen because those ends won't make them happy.

sense of good-for, in which something is good for you if it enables you to function as the kind of thing that you are. If the thing in question is a sentient creature, part of the way the creature functions is that she has inclinations, which track the things that are good-for her in the functional sense. In the case of the other animals, I can put this more simply: their inclinations track the things that enable them to survive and reproduce, and so to function well as the kinds of creatures they are. Once the inclinations are in place, the things that are good-for a creature in the functional sense are also ends for her, since she desires them for their own sakes. In that sense, her own well-functioning and the things that sustain and promote it constitute her final good.

To adapt the theory to the human case, we need an account of what it is for a human being to be well-functioning. I didn't talk about this much in *Fellow Creatures*, because I wanted to keep the focus on animals. I believe that the human good involves the successful formation of what I have elsewhere called a practical identity, which consists of the roles and relationships that make our lives worth living and our actions worth undertaking in our own eyes. There are therefore two aspects to the human good—we must function well as human beings, and as a part of that we must function well as the particular people we are—that is, as the bearers of the particular practical identities that we construct for ourselves. Furthermore, I think it is up to us how we balance these two aspects when they come into conflict. Something that is best for you as the bearer of a particular practical identity might not be best for you considered

simply as a human being, as when a role or relationship requires risks or sacrifices of your own basic human needs. This means that the construction of our good is substantially up to us—not just in the epistemic sense, that we know best what will make us happy, but in a practical sense, that we decide who to be, and therefore what counts as well-functioning for us. This is why I think that beneficence towards humans takes the form of respect for their autonomy. I don't think this leads to any collapse.

2. Now to Gary's second question. Gary thinks that we might need to talk about three "tiers" of beings, since I am running together rational capacity with moral capacity—as Kant sometimes, but not always, seems to do. That is, I am supposing that rationality in general is inseparable from moral rationality. Gary suggests that psychopaths show us a middle possibility, between humans and animals, since they are capable of reasoning in some sense. They seem to be what in the book I call "descriptively rational"—they act on considerations, and can tell you what they are—and Gary thinks we ought to respect their choices, so we owe them respect for autonomy, and not just beneficence. In that way, they are not like animals. But they are not moral, so they are not in relations of reciprocal lawmaking with us, so in that way, they are not like other human beings.

I feel somewhat awkward entering into a discussion with Gary Watson about psychopaths, since, I suspect like many of you here, I know everything that I know

about psychopaths from reading Gary Watson. Gary tells us that psychopaths characteristically suffer from prudential defects as well as moral ones—they are not very good at taking care of themselves and carrying out their plans—and Watson himself has written in defense of the theory that these conditions are linked. After describing these defects, Gary says:

> In failing to take their ends seriously, these subjects (i.e. pyschopaths) do not, after all, treat themselves as worth much. (Regarding your life as worthwhile entails valuing yourself.) An individual's controlling sense of the importance of how her life goes supplies a critical standpoint from which she might find compelling reasons for persisting in the face of distractions, for achieving coherence of aims, for self-correction and for regretting the way in which she has conducted herself.[3]

Gary argues that psychopaths seem to *have* the kind of reflective distance from their impulses that I claimed is characteristic of rational beings in the *Sources of Normativity*, but that in their case the reflective distance is, in Gary's words, "practically idle." The linkage between their moral and their prudential defects would then be between their inability to call their impulses into question from a critical standpoint supplied by their

[3] Watson "Psychopathic Agency and Prudential Defects"

own overall good, and their inability to call their impulses into question from a critical standpoint supplied by thoughts about the good of others.[4]

I am even more convinced than Kant was that instrumental and prudential rationality are inseparable from moral rationality. In "The Normativity of Instrumental Reason," I argued that instrumental rationality cannot exist apart from having some normative standard governing the choice of ends, and so from something like morality. Only an end judged good can give rise to a *reason* to take the means. In "The Myth of Egoism" I argued that if there is such a thing as prudence, it must depend on a principle that functions categorically, not on an instrumental principle. Watson's defense of the idea that prudential and moral defects are linked enables me to say what I want to about this case, which is psychopaths are not merely differently rational, but rather defectively rational.

To explain: There is a section of *Fellow Creatures* in which I argue against the argument from marginal cases. This is obviously not because I wish to defend the idea that only human beings count morally, but because I think the argument is based on bad

[4] In *Fellow Creatures*, I claimed that the application of the concept of the good requires rational empathy: to talk or think about what is good for another creature, you have to see the world through that creature's eyes. We might also think that prudence, in the sense in which it involves taking your own future good into account, involves something like empathy with your later self. Perhaps an inability to imagine *yourself* suffering from the disabilities of old age is at work in much youthful imprudence.

metaphysics. The argument from marginal cases claims that it is "speciesism" to accord moral standing to a human being who, as proponents of the argument often say, "lacks rationality" but not to, say, a dog who also "lacks rationality." Talking of "rationality" as if it were a property one might simply have or lack is part of the problem here. But the main problem with the argument is that to be a member of a species is not just to have a certain set of properties, but to function in a certain way. To function in a certain way is to exhibit a certain kind of unity – in a functionally organized object, everything depends on everything else, and a deficit somewhere in the system can throw off the whole. The reason we treat a human being who lacks "rationality," or rather some feature or aspect of rationality, differently than we treat an animal who also "lacks rationality" is that the human being is unable to function well or at all without that property, while a non-rational animal functions perfectly well without that property, since he functions in a different way altogether. There is a fairly clearly sense in which the defective human ought to function rationally, and our treatment of him should reflect that fact.

Now I think that the fact that psychopaths lack the critical perspective that is usually part and parcel of reflective distance places psychopaths squarely in the category of defectively rational beings. In fact, it makes them a good case for my argument. So I would say they are not on a second tier but are defective members of the "top" tier. What difference does that make? Watson thinks we would amputate the leg of a

wounded dog to save its life without regard to the dog's own wishes, if the dog even can have any wishes on such a subject, but we should not do this to a psychopath who objects. One thing we do in cases where a person lacks autonomy or is unable to access it (like someone in a coma) is try to do what they *would* choose if they could. For similar reasons, we might do what we know the psychopath prefers even if we think he isn't exactly autonomous. I'm not sure about that. But in any case, I agree that we should not treat the psychopath the way we treat an animal, even though I don't think the psychopath occupies a middle tier.

I should also say that I think there is such a thing as instrumentally practical *thinking* that falls short of instrumental *reasoning*. Many primates and corvids clearly can figure out and be motivated by means/ends connections, in some cases rather complicated ones. For example, there are crows who will fashion a tool in order to reach another tool that they need to get at something they want, so their thinking involves a grasp of a complex chain of causal connections. To see the force of an instrumental reason, though, is not just to be motivated take the means to the end, but to be committed to taking it unless certain kinds of new motivational factors appear on the scene—but *only* certain kinds, not just any.

I discussed this in another context in Fellow Creatures,[5] but let me adapt a simpler example I used in *Self-Constitution* to try to show you what I mean. If I decide I have an instrumental reason to go to the dentist tomorrow in order to get my toothache fixed, I nevertheless may rationally decide not to go to the dentist tomorrow if, when tomorrow comes, I have to rush to the help of a friend in trouble. And I know that when I make the decision—instrumental reasons are always conditional in this way. But supposed I am prepared to change my mind about going to the dentist because of any *new* motivational factor whatever—for instance, I think it would be okay to abandon my trip to the dentist because when tomorrow comes, I am more afraid of the drill than the toothache. Then I am not acting on an instrumental reason; at most, if I do go to the dentist after all, I am being motivated by a causal connection between an act and something I want. Deciding you have a reason to do something involves having a commitment to getting it done, unless something with the status of an *opposing reason* comes up.

Watson discusses a psychopath named Dever who decides to commit a robbery in Florida. Watson says,

> This activity is an operation of practical rationality, for it is the acceptance of an objective that to some extent structures and guides

[5] 8.5.2., pp.142-143

> his thought about what to do next (or soon). It now makes sense to him to take the evening train to Miami (rather than the subway to Brooklyn) and to get his hands on some burglary tools for the job.… But his aims do not provide a normative standpoint in the sense we are concerned with, for they provide no basis for self-criticism or evaluation of the objectives themselves. Dever may just lose interest in his plan, in which case it will no longer make sense for him to proceed to Miami. But he cannot see that fact as any kind of cause for self-rebuke.

This is exactly the sort of case in which I would have said merely instrumental practical thinking, not instrumental reason, is involved. Of course you might say that I am just using "reason" and "reasoning" in a more specific sense than people usually do. But that doesn't matter to my point. However we want to use the word "reason," it looks as if what Dever does may differ only in degree, not in kind, from what a crow does who makes a tool in order to reach another tool does. So maybe we should say he is not rational after all. He is just more intelligent than a crow. I don't quite think that, because Dever acts on considerations he's aware of and can think about. I think it is possible that Dever regards the considerations on which he acts merely as explaining what he does, not justifying it, and that he wouldn't understand the distinction.[6]

[6] If that's right, his attitude is weirdly impersonal. But it also suggests you can reason theoretically without being able to reason practically, and since I think those are essentially connected as well, I would need a further story about how his defect makes that possible. Harder to see how that might go.

3. I now turn to Hilary's questions. First she wonders if we really confer absolute value —goodness that gives a reason to everyone—on our ends whenever we make a choice. Suppose, for example, you reach for a Diet Coke. Does that confer absolute value on your having a Diet Coke? Of course when we are trying to explain the way in which this act gives rise to reasons for others, it might seem more plausible if we say that thing that has absolute value is my quenching my thirst in a way that will give me innocent pleasure, or something like that. But Hilary asks, why not just talk in terms of rights: I take myself to have the right to determine my own conduct so long as I don't interfere with anyone else's rights.

There are certainly cases in which we don't feel we have to justify our conduct to others, in the sense that we don't owe them an explanation of, say, why we choose a Diet Coke rather than a Pepsi or a glass of almond milk. Telling someone who presses such a question that you have a perfect right to choose your beverage for yourself seems in order. But you don't tell yourself that you drink a Diet Coke because you have a right to: after all, you have a right to choose all sorts of beverages, and that thought doesn't sort among them, so that doesn't exhaust your reason.

Kant says we always act for the sake of good, but I should also say that I think the thesis that we always act, as people say now, under the guise of the good is true and important independently of Kantian arguments. I believe that when we act we monitor

our movements in ways that are controlled and regulated by the value we set on our ends, and we make decisions that would be impossible to make if we thought of our ends as having no value at all.[7] Perhaps the value you assign to having a Diet Coke makes it worth reaching out your hand for one, but should unexpected obstacles appear in your path, you have to know *how much* it is worth in order to decide what to do about them. Do you have to get up out of your chair to get it after all? Do you have to knock something out of your path? Was someone else reaching for the very same Coke at the very same moment? Ends that have no value do not have an amount of value, unless you take that amount always to be zero, in which case it is unclear why it was even worth the effort of reaching out your hand.

Now I realize you might think that the demands of the guise of the good thesis are met by something's having a value that is not absolute, but just consists of how good it is *for you*; you only need to know how much value the Coke has for *you* in order to make these decisions. So you might think that the notion of good-for will serve these purposes, and we don't need absolute good. I think that is wrong when we are in interaction with others. Of course you might think that all we do is set a value on our ends based on how good they are for us, and then just duke it out with other people when it's worth it; or, more nicely, you might think we set a value on our ends, based

[7] I think the other animals do this too, which is one of the reasons I think they see the world in a valenced way. See "Prospects for a Naturalistic Explanation of the Good."

on how good they are for us, but when we enter into interaction with others it becomes a question of rights. I don't think this is an adequate way to think of what goes on in interaction, but I don't have a worked-out argument to give you on that point now.

Hilary thinks we sometimes act without any end, and cites the example of someone who tells a friend that her partner is cheating on her simply because she thinks it is the right thing to do. Kant asserts that we always act for the sake of an end, although it can be an end that must be respected, not one that we are trying to promote, and he thinks that is true of any action, including moral action. But actually, this doesn't matter for my purposes here. I invoked the claim that we treat our ends as absolutely valuable and therefore show that we regard ourselves as ends in ourselves specifically as part of an argument that, at least as I read it, turns on a question about why we regard it as rational to do merely permissible actions, and why we think that our doing so makes a claim on others—at least not to interfere with us, but possibly to help us as well. Even if moral actions don't have an end, that doesn't matter for my argument unless you think moral actions are the only ones we do.

6. Hilary thinks that I should not say that we regard our ends as good absolutely; we only regard them as prima facie good, and give them up when they are incompatible with the goods and claims of others. Gary, in a footnote in his commentary [which he

didn't read out], voices what I think is in part a similar point—that it is obvious nothing can have absolute value in my sense.[8]

Hilary thinks that if I would just scale my claim back to the claim that our ends have *prima facie* goodness, I could avoid what she claims I identify "as a serious problem with my account —namely that there is a "rift in the good." I do claim that there is such a rift. The rift arises from the conflict between these claims: the claim that the good of each creature (that is, person or animal) is good absolutely, the claim that what is good absolutely must be good from the point of view of every creature—that is, it must be at least consistent with that creature's good—and the plain fact that the goods of different animals are at odd with each other in a system of nature which depends on competition for scarce resources and the relation of predator and prey.

I want to make it absolutely clear that I do *not* regard this as a serious problem for my account. I think that all of these inconsistent claims are true, and that that is a serious problem with the natural *world*. The natural world resists the consistent application of

[8] At some points in the book I say that absolute value is value for everyone, or from anyone's point of view, including the point of view of animals. Gary protests that "it seems too obvious that nothing has absolute value in this sense." His complaint is that it just doesn't matter to most animals whether I get what is good for me. But he concludes that "the most that Korsgaard can hope to show—and all that in the end she needs to show—is that the sort of importance rational agents give to their own ends commits them to valuing the ends of every sentient creature." Hilary would say this means all I can show is that we have to assign the ends of the other animals *prima facie* goodness.

what I take to be *our* concept of the absolute good—a state of affairs that is good for all creatures who have a good. As long as we confine moral standing to human beings, we can avoid facing up to the implications of this problem, both because, as Hilary says, a rational being's good cannot be furthered by acting wrongly, but also because human beings can for the most part adjust both our social systems and our own personal good in ways that make the good of all people compatible with each other. But we can't construct a state of affairs that is good for everyone once we admit that animals have moral standing. I do not think that we should domesticate our moral concepts so that they fit neatly onto the situation, either by backing up to the notion of *prima facie* good, as Hilary proposes, or by denying that what happens to animals matters very much, as most people do. I think we should face the fact that we live in a world that is hostile to the application of human moral concepts, and also that we should stop using this fact as an excuse for treating the non-human animals with the gross brutality that we do. We should do the best that we can to make the world a good place for all of its inhabitants, even knowing that we cannot succeed, because nothing short of that is adequate to our concept of the good.

Appendix: Hilary's Parenthetical Questions

1. Hilary may be right to argue that there is more to be said for the idea of treating someone else morally because you owe it to yourself rather than because you owe it to them. I am not sure about her example—refraining from torturing a terrorist because you owe it to yourself rather than because you it to the terrorist. She cites Kant's view that lying violates a duty to yourself even if the person you lie to has forfeited any right to expect the truth. I think that lying is treating a person as a mere means in the plainest way—you look right into his eyes as if you were addressing a fellow person, but you say whatever you think is necessary to get him to do what you want. In my paper "The Right to Lie: Kant on Dealing with Evil," I argued that this is sometimes justified, but always morally regrettable, and I do not see how to separate that last thought from the thought that you owed him something different considered simply as a fellow human being. I don't think Kant's claim that we annihilate our own dignity by doing this is helpful, since I don't see how we can establish that it does annihilate our dignity without first establishing that it is wrong. Hilary points out that it is a further, and not a logically necessary thought, that the reason the lie wrongs oneself is that it harms one's character. I agree, and it was one of the points I made against Kant's argument in the book §6.2.3, p. 101. But in any case Hilary doesn't disagree with me about whether we have duties directly to animals.

2. Hilary wonders why I don't make more of Kant's contradiction in the will test; nearly everything I say in the book about the universal law formulation is about the contradiction in conception test. There are two reasons. First of all, I think the contradiction in the will test as presented in the *Groundwork* just doesn't work. Second, Kant thinks that the two contradiction tests give rise to different kinds of duties—the contradiction in conception test to perfect or strict duties (we can argue about which of those it gives rise to, since they are not exactly the same), and the contradiction in the will test gives rise to imperfect or broad duties. Elsewhere I have argued that the contradiction in conception test has trouble handling what I called "the problem of natural actions"—for example, the use of force and violence to achieve your ends. Universalizing natural actions does not undercut their efficacy for achieving their ends: for instance, no matter how many people kill their enemies, killing your enemies will still be an effective way of getting rid of them. Hilary proposes to treat these under the contradiction in the will test instead. She says, "a world in which people were forever killing people who got in their way, or beating them to a pulp, would not be a world that was hospitable to rational agency." But if we take Kant at his word, this would make the duty not to kill people, or beat them to a pulp, imperfect or broad duties. This just seems wrong to me. In any case, we know that a world in which people treat animals as if they were objects for our use works perfectly well for achieving our own ends, because that is the world in which we live.