

*Forthcoming in Oxford Studies in Metaethics Vol. 2*

**VALUE AND AUTONOMY IN KANTIAN ETHICS**

Robert N. Johnson

Kantian ethics can at times appear to defend the position that there is a unique sort of value that plays a foundational role in morality. For instance, Kant's most well known work in ethics, the *Groundwork of the Metaphysic of Morals*, begins by trying to establish that a good will is good 'without qualification' and then ends with a first statement of the fundamental principle that divides right from wrong, the Categorical Imperative.<sup>1</sup> This presentation can make it seem as if Kant believes the authority carried by the Categorical Imperative is somehow supposed to be grounded in the value of a good will. Again, the humanity formulation of the Categorical Imperative, the formulation that tells us we must respect the humanity in ourselves and others by treating it as an end in itself, appears to allude to a special value possessed by some feature of persons, their humanity, and then explain the authority of moral obligation by way of that value.<sup>2</sup> This extolling of the value of humanity and the dramatic refrain about the unique value of a good will both appear to portray Kant as telling us to notice the peculiar value that they possess and see that this value demands that we adjust our deliberation and actions. We appear to be told that the good will and humanity are bits of metaphysical glitter, jewels carrying their value around with them, and that this unique glitter is source of the authority of moral obligation.

---

Thanks to Peter Vallentyne, Ben Bradley, Jon Kvanvig, Richard Dean, Matt McGrath, those at the 2<sup>nd</sup> Annual Metaethics Conference at University of Wisconsin, Madison and at the University of Tennessee-Knoxville, and anonymous referees for helpful comments and suggestions.

<sup>1</sup> Kant (1996) 4:393-405.

<sup>2</sup> Ibid., 4:428-29.

I say that Kantian ethics can *appear* to say such things because I believe that it is in fact just an appearance. The value of humanity or of a good will does not in fact underwrite the authority of moral obligation. As I shall eventually argue, that authority must come solely from the fact that it is a demand of our own reason. If this last claim sounds familiar, it should; it is at the core of Kantian orthodoxy. But as familiar as the claim is, many have recently gone on record taking issue with this orthodoxy. While the orthodoxy holds that in fact Kantian ethics does not ground moral requirements in a value of some sort, in their view the value attributed to our humanity or to a good will is indeed the source of the authority of moral obligation. In what follows, I defend the orthodoxy against this attack. I argue that their position is in fact utterly incompatible with central doctrines of Kantian ethical theory, in particular, with the view that the human will is self-legislating or autonomous. The distinctive reason we have to conform to moral obligation, in Kantian ethics, stems simply from the autonomous nature of rational agency itself, not from its value or the value of anything else. Thus, if the human will is to be autonomous, then not even the value of a good will or of the humanity in persons can be the source of the authoritative reason Kantian ethics claims there is to conform to moral obligations. In short, there is no room and no need for metaphysical glitter in Kantian ethical theory.

In the first section, I detail the claim about value in Kant's ethics that a number of Kantians have recently defended. In the next, I explain why I think the Kantian doctrine of the autonomy of the will is incompatible with this claim. In the final section, I discuss how we should understand the relationship between Kantian views on value and the authority of moral obligation. Here, I bring elements of the Kantian theory of value to bear on the familiar discussion of the nature of the value of the good will and humanity.

### 1. *The Anti-Deontological Kantians*

What I believe to be incompatible with the doctrine of the autonomy of the will is the claim that an appeal to some (however unique) value is required to explain the authority of moral obligation. By the ‘authority’ of moral obligation, I mean both the reason there is to comply with our obligations and the particular force or standing that this reason has. Thus, the claim goes, in order to explain how there could be a reason to conform to obligations that can often require significant personal sacrifice, we must appeal to some special value that is realized in or achieved by these actions. This is the way teleological views such as utilitarianism explain the authority of moral obligation. In the case of utilitarianism, that explanation is in terms of the nature of what she claims is of value, the general happiness. The reason to conform to moral obligations, in her view, is that such actions promote the overall good, which she identifies as general happiness. According to such a view, there is a good, achieving the general happiness, which provides reason to comply with moral obligations. And the fact that this reason is found in promoting the general good, and not just one’s own, is the source of its special status. Likewise, some think that there must be a value provided by Kantian ethics. To be sure, the Kantians who believe that there is some sort of grounding value in Kant’s ethics do not think of that value as a utilitarian does, as an *outcome* of conformity to moral obligation. Instead, they simply believe that some good is realized by, or is in some other way connected to, performing one’s obligations. And that value, they say, is contained in the value possessed by our humanity, our freedom, or the good will.

As a first example of such a view, consider Paul Guyer's position. He holds that for Kant "freedom has an 'inner value, i.e., dignity'" and that this "is the fundamental normative fact" that is "the premise on which Kant's" moral philosophy rests.<sup>3</sup> Freedom gives us such an extraordinary sort of dignity, a dignity which so overwhelms the perceived value of satisfying any particular needs or inclinations that it can be immediately recognized to fulfill the expectations of unconditional value raised by our ordinary conceptions of morality and duty.<sup>4</sup>

Thus, Guyer's position is that Kant's moral philosophy finds a special value, dignity, possessed by the freedom of a rational will, and that this value is the source of the unique reason to comply with the Categorical Imperative. For that value "overwhelms the perceived value" of satisfying our own interests, thus showing that there is some good in complying, albeit perhaps not our own personal good. That, in Guyer's view, is the way in which we explain the authority of moral obligation.

Barbara Herman also defends a Kantian position that appeals to a distinctive value. She argues that "Kant's project in ethics is to provide a correct analysis of 'the Good' understood as the ultimate determining ground of all action" and so it is false that his view implies that "no moral principles subtended from a concept of value can explain obligation".<sup>5</sup> She defends the view that the principles of practical reason to which a good will conforms *themselves* constitute "a conception of value". Her defense is meant to counter the criticism that there is no such value in Kantian ethics. The criticism turns on the idea that unless there is a value that provides reason to conform to moral obligation, "the rationale for

---

<sup>3</sup> Guyer (1998), p. 33.

<sup>4</sup> Ibid., p. 28

<sup>5</sup> Herman (1993), pp. 209-10.

moral constraint is a mystery”.<sup>6</sup> Thus, Herman as well holds that a unique value, the practical principles characterizing ‘the good will’, provides a reason to conform to moral obligation in Kantian ethics.

Another Kantian, Allen Wood, interprets Kant as holding that the value of humanity provides such a reason. He argues first that our own rational nature is itself “the source of the fact of [other things’] goodness—indeed of the fact that anything at all is objectively good.”<sup>7</sup> The value of our own nature is the source of reason-giving value in much the way that the source of the authority of someone’s recommendations to us is our respect for that person’s authority on the question. The value of the recommendations depends on the value of her authority: advice is only as good as the expertise of the advisor. That means that the choices of a rational will “can confer objective value on other things only if it is presupposed that it has objective value”.<sup>8</sup> Hence,

if rational nature is in this way the prescriptive source of all objective goodness, then it must be the most fundamental object of respect or esteem, since if it is not respected as objectively good, then nothing else can be treated as objectively good...rational nature is presented as the only thing that could answer to the concept of an objective end or an end in itself.<sup>9</sup>

Wood’s position is that while every other value is derived from the fact that they are chosen or pursued by a rational will, the rational will itself must have a unique intrinsic value, a value not derived from anything else. Our rational nature, our humanity, is a self-standing

---

<sup>6</sup> Ibid, p. 210.

<sup>7</sup> Wood (1999), p. 130; see also pp. 157-8.

<sup>8</sup> Ibid, p. 130

<sup>9</sup> Ibid.

value. And that self-standing value is the source of the universally binding reasons to comply with moral obligations.

Wood develops his position in response to Christine Korsgaard's reconstruction of Kant's argument for the value of humanity. Korsgaard herself appears to be of two minds on this issue. In defense of the Kantian view of value, she argues, as does Wood, that

good things are good in the way that Ross describes as relational, because of attitudes taken up towards them or because of other physical or psychological conditions that make them important to us. Only one thing – the good will itself – is assigned an intrinsic value or inner worth, and even the argument for that is not ontological. If we regard ourselves as having the power to justify our ends, the argument says, we must regard ourselves as having an inner worth – and we must treat others who can also place value on their ends in virtue of their humanity as having the same inner worth.<sup>10</sup>

She thus regards the good will as having a non-relational value, a value that is the source of our capacity to justify our ends. This position seems to be in line with Wood and the rest, the position that the unconditional value of the good will is a grounding source of reasons for moral demands.

But Korsgaard has also defended a view she dubs 'procedural realism', which appears to reject this position.<sup>11</sup> Procedural realism denies that there are any substantive normative entities or properties to which our practical terms refer. Those who assert that there are such entities she dubs 'substantive realists'. The problem with substantive realism, she argues, is that "the appeal to the existence of objective values cannot be used to support our

---

<sup>10</sup> Korsgaard (1998), pp. 272-3.

<sup>11</sup> For an excellent discussion of Korsgaard's metaethical views, see Hussain and Shah (forthcoming).

confidence” that we have moral obligations, since it is that very confidence that supports our conviction that such values exist.<sup>12</sup> The only genuine grounds for asserting the existence of objective values is just this confidence, and in her view we have this confidence quite apart from proving their existence. An obvious way of construing this position that objective values cannot be used to support our confidence in moral obligation is as a denial of the claim that it is the objective value of humanity that is the source of the authority of moral obligation. That is, Korsgaard appears to mean that no objective values are sources of reasons to comply with moral obligations.

Another reason for thinking that Korsgaard would deny that a value grounds the authority of obligation is her view that “normative concepts like right, good, obligation, reason are our names for the solutions to normative problems.”<sup>13</sup> I take it what she means to say is that value terms do not apply to value entities, in particular, value entities that could generate reasons for action. Of course, normative problems clearly do have solutions that refer to things that appear to give us reasons for actions. ‘What ought we to do for the homeless?’ is a normative problem with a solution – the offers of food, shelter, and medical aid, for instance — that are valuable. And it seems their value is what gives us a reason to make these offers. But, for Korsgaard, in fact there is no substantive value, such as the value of an offer of food, shelter and medical aid, which generates a reason to do what we ought and make that offer. If anything would be the source of a reason on this view, it would be that rational agents are disposed to agree on giving such an offer to the homeless.

While these positions seem to support the view that Korsgaard holds that no value is the source of reasons to conform to moral obligations, they fall short of explaining

---

<sup>12</sup> Korsgaard (1996) p. 40

<sup>13</sup> Ibid, p. 47. I assume that she didn’t mean to say that concepts are names.

everything that needs explaining. It is in how this further part of the explanation is filled out that will answer the question whether she accepts the view that I've attributed to Guyer, Wood and Herman. In fact, what is left to explain is what Wood's view adds to Korsgaard's, namely, the explanation for the fact that we often decide what to do by appealing to values we regard as in some way supporting our actions, and the fact that, when reasoning together, take ourselves to be trying to converge on values that will provide reasons for acting. What justifies following the rules or procedures of reasoning that produce this convergence if it is not the value of the rational agency of the deliberators who converge on them? One could say that the reasons supporting a choice flow from the value of what that choice aims at, and procedures would be justified only if they yield such values. But value facts or properties that could explain the rationality of such choices are the very things that Korsgaard denies exist. So we are left to explain the reason we have to comply with moral obligation as in some way coming from the value of humanity in ourselves and in others.<sup>14</sup> Korsgaard does not explicitly say so. But, often enough, she treats certain values as being exceptions, saying, for instance, that some things, a good will for instance, "must obviously carry its own value with it".<sup>15</sup> Although the evidence is equivocal, I shall assume that she would hold, as do the others, that while the value of everything else is nothing more than its possessing a relational property, of being the object of a rational will or being the point of agreement between deliberating rational agents, the good will and humanity possess values that are "carried around with" them, and that she thinks these values provide reasons to comply with moral obligations.

---

<sup>14</sup> See also Korsgaard (1997), p. 218.

<sup>15</sup> Korsgaard (1998), p. 257.

Thomas Nagel will be my final example of a philosopher who holds that Kant's ethics views some unique value to be the source of reasons. Kant's ethics, in his view, implies that "freedom can be pursued and approached only through the achievement of objective and ultimately ethical values of some kind".<sup>16</sup> The 'achievement' of such values does this by allowing "the will to expand at least some of the way along the path of transcendence possible for the understanding" and attain a practical kind of 'view from nowhere'.<sup>17</sup> The will expands to this view, not by attaining the sort of 'external' view available from the point of view of the sciences. That sort of external point of view is inconsistent with action, which always requires to some extent taking one's own internal point of view as agent. Yet free willing is willing not merely what is of value to you, from your own point of view, but what is of value from an objective point of view. So being guided by objective values is supposed to provide whatever degree of freedom we think we enjoy in action, by providing a certain sort of reason for action, a reason that supports conforming to deontological constraints. This well-known line of reasoning from Nagel's work seems to me to show that he too defends the claim, in the name of Kantian ethics, that some objective value is the source of the authority of moral obligation.

### *(3) Autonomy and Reasons*

So far, I have outlined a trend I see among Kantians to portray value, such as the value of freedom, humanity or the good will, as a source of a distinctive reason to conform to moral obligations. In this section, I explain why I think that they must, to remain consistent,

---

<sup>16</sup> Nagel (1986), p. 137

<sup>17</sup> Ibid, p. 136.

deny that the value of these things provides such a reason. In the following section I will go on to argue that this is nevertheless consistent with holding that their value is unique.

It is instructive to contrast Kant's views, at least as I portray them, with that of others who also deny that some objective value gives us a reason to conform to moral obligation. These others reject this position simply because they think something subjective provides such a reason. For them, it is, for instance, satisfying our desires, not realizing some objective value, which constitutes whatever reason there is to comply with moral obligations. The Kantian view is different from such desire or interest-based explanations of the authority of obligation in holding that what we call 'good' or 'valuable' is whatever is an object of *practical reason*, and it is our own practical reason that is the source of the authority of obligation, not our desires nor the value of the objects of practical reason. According to this view, the 'practical' in 'practical reason' concerns the will conceived of as a faculty consisting of something distinct from the sorts of empirical psychological states collected under the term 'desire'.<sup>18</sup> Details about the faculty of practical reason are not important. The key idea, deployed the Kantians I have been discussing, is that things, or at least most things, are valuable because they are objects of rational (in a Kantian sense) choice. What is at issue is whether, in addition to things that acquire value by being related in the right way to rational choice, there is another sort of value, the value supposedly residing in the good will or humanity, that does not get its value in this way, is the condition or source of the value of the objects of rational choice, and is ultimately what provides a reason for conforming to moral obligation.

In my view, Kantian ethics must deny that there is such a value. It must deny this because it is inconsistent with claims that are central to Kantianism, namely, that a) there are

---

<sup>18</sup> Kant (1996) 5:60

requirements that are absolutely binding on rational agents, b) there are such requirements only if rational agents are wholly self-legislating or are autonomous, and c) rational agents are wholly self-legislating only if no value explains their authority over us. That a) is a core Kantian doctrine should be familiar enough. This claim just comes to the idea that there are categorical imperatives, or commands that apply to us no matter what our personal contingent reasons might be for or against complying. Admittedly, the existence of such requirements is controversial. That they in turn require the autonomy of rational agency, as b) states, is more controversial, though still a familiar aspect of Kantian ethical theory.<sup>19</sup> It is c) that most concerns the relationship between autonomy and value.

First, let me give you an intuitive sense of why autonomy commits Kantian ethics to denying that value is a source of reasons. Consider a parallel example: the Divine Command theory's resolution of Euthyphro-style dilemma. That dilemma begins with the assertion that God loves (or responds in some appropriate way to) all and only good things. This raises the question, Why? Is it because their value provides a reason for God to give them their due, His love? But if God loves a good thing only because its goodness gives him a reason love it, then its goodness explains the appropriateness of God's love for it, and this is incompatible with God's omnipotence. The value that provides a reason for God to love it would be a constraint on God's love in the sense that God must respond to the reasons provided by the value of things or else fail to have the requisite response. The alternative is to say that those things are good because of God's love for them. God's love explains their value. But then goodness looks like an arbitrarily distributed shadow cast by God's attention. For if God loved some entirely different set of things, then those other things would have been good.

---

<sup>19</sup>This has been discussed under the heading of 'the Reciprocity Thesis' recently by Henry Alison in his (1986).

The Divine Command theorist opts for the second horn, and then is saddled with the problem of explaining why goodness isn't arbitrarily distributed after all.

Kantians must resolve a similar Euthyphro-style dilemma in the same way as the Divine Command theory. What possess value on the Kantian view are all and only the objects of rational agency. Now if value is the source of the reasons for the pursuits of rational agents, then the authority governing rational agency is external to that agency itself, in the value of the things that are its objects. But on the Kantian view, rational agency must be autonomous, in the sense that the requirements binding it are wholly self-generated and self-imposed. The autonomy of reason, the central guiding idea behind Kantian moral theory, is thus the very foundation of the case against the claim that there is some value that provides reason to conform to moral obligation. Autonomy requires that value not be a source of reasons.

Rational agency must be autonomous because, as I mentioned above, if there is to be an absolutely authoritative rational requirement such as the Categorical Imperative, then it must be self-legislated and gain its authority from this self-legislation. The line of reasoning is this: A practical requirement is binding only if the agent can voluntarily comply with it. But an agent can comply with a requirement only if there are reasons for her to comply. Now if an agent is bound absolutely to comply with a requirement, then she must always comply with it. But then if she must always comply, there must always be a reason for her to comply. But for any requirement, there will not always be personal reasons for an agent to comply – reasons that come from her own contingent circumstances. In a circumstance in which there are no contingent personal reasons, then, there must be some other reasons. But the only reasons that will always be present for complying with an absolutely binding requirement will just be whatever reasons there were for imposing that absolutely binding requirement on the agent in the first place. So, if the agent is bound absolutely to a requirement, then, given that

the ability to comply requires reasons, it must be possible for her reasons to be those on the basis of which the requirement was imposed on her (or legislated).

Now if it is possible for her reasons to be those on the basis of which the requirement was legislated, then she must be able to appreciate fully why the absolutely binding requirement came to be legislated. And if she is able to appreciate this, then she must be able to engage fully in the deliberation that led to its legislation. But she will be able to engage in that reasoning herself only if she herself possesses whatever capacities and hence authority a legislator of such a requirement can claim that is sufficient to generate and enact it. So an agent who is bound absolutely by a requirement must be no different from its legislator, the source of its authority over her.<sup>20</sup>

A rational agent will therefore be wholly self-legislating only if the authority of the principles governing her agency comes from the fact that she herself is the author of those principles. The reason she must conform to these principles is thus that she gave them to herself – not because some good comes from or is realized by following them. And rational agents give themselves the laws that they do because it is in the nature of their rationality to do so. As it happens, on Kant's own view, rational agency operates on the basis of laws valid for all rational agents, and it is essential to (and not *analytic of*) rational agency that it do so. This is because rational agency is essentially a kind of causation. Since any causation in Kant's view brings with it universal laws, it is essential to rational agency that it govern itself by *universal law*, laws valid for all rational agents. This is how Kant proposes to show that being governed by the Categorical Imperative is essential to rational agency. My argument, however, does not require this last step connecting autonomy with the universal law version of the Categorical Imperative through the idea of causation. What autonomy of the will

---

<sup>20</sup>I here draw on Andrews Reath's reconstruction of this argument in his (1994).

requires is only that the explanation of the authority of the principles governing the will comes from the fact that the will is the source of those principles. And if the reason for you to conform to a law is the fact that you gave that law to yourself, then the reason does not derive from any value, such as the value of your will or your humanity. To be sure, your humanity is, in Kantian ethics, composed in part of your capacity to lay down practical laws for yourself. And that humanity is of unique value. But the reason you must conform to moral requirements does not derive from the unique value of this capacity, in you or in anyone else. It derives from your exercise of that capacity alone. So the thesis that the will is autonomous is not compatible with the claim that the source of the reason there are to conform to moral obligation is a value of some sort.

There is, then, no room for the metaphysical glitter of a special value in Kantian ethics. But isn't there yet a need for it? For if no value is realized, achieved, respected or otherwise brought about by conforming to moral obligations, then what reason is there to do so? Kantian ethics, one might object, is satisfied giving no reason at all. This objection is right insofar as Kantian ethics does not appeal to the value of something to give such a reason. But the assumption of the criticism is that a reason requires something of value, and the Kantian position I have laid out above contests just this assumption. There is a reason to conform to moral obligations, but that reason is that you demand it of yourself. Kantian ethics at bottom the idea that the reason you should conform to moral requirements is the fact that you imposed them on yourself, and you do this simply because you are a rational agent. Quite apart from whether the fundamental moral principle is the Categorical Imperative, it must, according to an ethical theory in which rational agents are autonomous, be constitutive of rational agency that the authority of its principles be grounded simply in

the fact that it is their author. And that, in turn, means that the fact that you gave yourself this principle is the only reason that exists in every circumstance to conform to it.

To be sure, the autonomy of rational agency remains a controversial aspect of Kantian ethical theory. My point is that *given* Kantians are committed to it, they simply must reject the idea of an authoritative good – as it might feature in the value of a good will or the value of humanity as an end in itself – a value that explains the authority of moral requirements. But, as I hinted at the outset, if, as I think, Kantians are committed to rejecting an authoritative good of any kind, this raises questions about views that also seem central to Kantian ethics, in particular, views about the value of the good will and humanity. How can humanity and the good will possess the unique sorts of value that they supposedly do yet not provide reason to conform to moral obligations, such as the demand that we respect humanity in ourselves and others? How could its value *not* be the reason we are to treat humanity with respect? I will turn to this last question in the next section to explain how I think they should be answered.

#### 4. *The Value of a Good Will and Humanity*

Let's return for a moment to Wood's argument for the value of humanity and its centrality to explaining the authority of moral obligation. Wood's view is that since rational choice is the source of all value it must be objectively good. This, for instance, is the sort of reasoning Kant seems to engage in here:

Nothing can have a value other than that determined for it by the law. But the law-making which determines all value must for this reason have a dignity—that is, an unconditioned and incomparable worth—for the appreciation of which, as necessarily given by a rational being, the word '*reverence*' is the only becoming expression. G 4:436

One can easily read this passage as asking us to accept the claim that humanity has a special value, and it is just this value that is the source of the authority of morality. Indeed, Kant appears to be saying that there is a difference between values, those that are authoritative and those that are not. The dignity and worth possessed by a rational will, he appears to say, is quite unlike other sorts of value, having ‘incomparable worth’. That incomparable worth explains the value of every other thing, and because of this, deserves our reverence. And that it deserves reverence is the reason to treat it in the various ways demanded by the Categorical Imperative. Wood appears to be reading such passages from Kant in this way, as saying, in summary, that the authoritative value of humanity is the source of reasons to conform to moral obligations.

Nevertheless, I think that this tempting reading is entirely wrong. We should read such passages about the value of humanity and the value of a good will in a different way. We should, in other words, take Kant at his word when he says “nothing can have a value other than that determined for it by the law”. “Nothing” means just that, nothing. So this must apply to “the lawmaking that determines all value” as well everything else. If it – that is, rational willing – is of value, it is because it too is an object of a rational will. This sounds dark and tautological, but it is not. Let me explain.

Assume that to possess a good will is, very roughly, to be deeply committed to the principle designated as Categorical Imperative. By ‘commitment’ I mean a disposition to affirm the Categorical Imperative as the ultimate standard for one’s behavior (although not formulated necessarily in the precise words used by Kant himself), to try (absent irrationality) to conform to it and to disapprove of oneself and others when deliberately or negligently failing to do so. Assume further that a disposition to be committed to this principle is that element of our humanity that gives humanity its special status in Kantian ethics. This would

be a higher-order disposition, a disposition to acquire the dispositions I have said characterize the commitment that is a good will. That second-order disposition, given it is characteristic of every person's will, is present in every circumstance in which we act. So the value of humanity will be the value of having this second-order disposition to acquire a commitment to the Categorical Imperative, and the value of the good will is the value of the realization of this second-order disposition.<sup>21</sup> I shall ignore for present purposes any differences there might be between, on the one hand, humanity and a good will, and on the other, freedom, 'rational nature' and other things regarded as of special value by the Kantians I've so far discussed.

Recall that the official Kantian theory of value *in general* (that is, without specifying the kind or modification of value, such as moral or nonmoral value) is that to be good is to be an object of a rational will, where 'rational' is understood in terms of the Kantian theory of rationality.<sup>22</sup> In short, what has value is whatever is an object of a rational choice. The Kantian views I have been discussing so far hold that the value of every good thing *except* the values of a good will and our humanity comes from being their objects of rational choice. I believe that this exception is a mistake. The value of the good will and humanity also comes from their being the objects of rational choice, and not the other way around. To understand how this is so, we must discuss the variety of ways in which something can be an object of choice.

There are two quite different kinds of objects of choice for Kantians, means and ends, as is implied by the Kantian slogan 'who wills the end wills the means'. We choose our means and, according to Kantians, our ends, making each a distinct kind of object of choice.

---

<sup>21</sup>Thanks to Judy Thomson for pressing me to clarify this.

<sup>22</sup>Kant (1996) 5:57-66.

Further, something's being an end of one's choice does not preclude its being a means in another and vice versa. My means of attaining the end of getting ice cream was a trip to the parlor, but getting the ice cream might in turn have been a means of satisfying my hunger. Moreover, something's being an end of one's choice does to preclude it from being a means in that selfsame choice. If my end it to engage in some activity, such as golfing, choosing to golf makes golf both the end I'm aiming at in my choice and the means of fulfilling it. Finally, not only, according to the Kantian view, *can* our ends be objects of choice, but they *must* be, since every object of will must contain an end, and all of a rational agent's ends must be chosen.<sup>23</sup>

When means are objects of the will, their value is only conditional and extrinsic – they are good only on the condition that the ends they serve are objects of our rational choice. It is the property of realizing or producing those ends that make them the objects of our choice. Humanity, however, is supposed to be 'an end in itself', never to be treated as a mere means to our personal ends, and this is supposed to mean that it is intrinsically valuable. Now there are different senses in which something can be an end. In one sense, an end is simply whatever we choose to produce or bring about in the world by some means. For instance, if having some ice cream is my end, then having some ice cream is an event in the world I set myself to bring about. Adopting these kinds of ends guides actions in that when I set myself to pursue them, I then must find actions that will be means of producing them. Choosing or willing an end is in this way a source of a law for my action: Willing the end dictates that rational agents do something, namely, act to bring about that end.

---

<sup>23</sup>Kant (1996) 6:380, 385. I argued in Johnson (2002) that this doctrine is incompatible with another doctrine Kant holds, that we have our own happiness as our end by natural necessity (e.g., see Kant (1996), 4:415-16, 6:382).

Humanity is not an end in this sense, but a good will can be such an end. For we cannot choose to have the second-order disposition of the sort we have identified as our humanity; that is simply something we have in virtue of being rational agents of the sort that we are, not something we can produce. So humanity is not valuable because we rationally will to produce it. But a good will can be said to be an object of choice in this sense, since a good will is the realization of the second-order disposition of acquiring a commitment to the Categorical Imperative of the sort I described above. Such a thing can be brought about, indeed can only be brought about, through choosing to acquire that commitment. As such, the value of a good will comes from its being the object of a fully rational choice, as being the end of that choice. That is to say, a choice to acquire a good will is a choice that is completely determined by Kantian rational principles, including the Categorical Imperative. Thus, the good will can be an object of rational choice and this is all that its goodness amounts to.

However, being the product of a choice is not the only sense in which something can be an end. When I shop, one of my ends is to economize.<sup>24</sup> But this is not something I set myself to produce. It is better understood as something that *prevents* me from acting in a variety of ways. For instance, choosing to economize prevents me from buying name brands instead of generic brands, from buying the first item I see, and so on. Some things, that is, are ends in the sense that they are that *against which* I may not act while pursuing other ends. Humanity is clearly meant to be an end of this sort, and hence can be an object of choice in this sense. Humanity is an object of my choice in the sense that economizing is an object of

---

<sup>24</sup> I borrow this example and the explanation of this distinction in ends from Barbara Herman in Herman (1993), p. 14.

my choice: It is a limit on my other ends, and so is a ‘negative’ end in the way that economizing limits what I may purchase.

Economizing, of course, is not an end every rational being must have. Some make economizing an end, others do not, and although it often limits the brands I myself buy, my health is more important than economizing, limiting how much I will rein in my spending. But my rationality does not depend on whether I have chosen economizing as my end, or with whether economizing limits my health or vice-versa. By contrast, humanity is supposed to be an *objective* end, an end that all rational beings must have, no matter what other ends they have and must limit every other pursuit, while no other pursuit may limit it. Hence, it limits what I may do – if I am to be fully rational – when I pursue my positive *and* subjective negative ends.

Although it is not something produced by my actions, humanity is also supposed to be in a different sense also a *positive* end. Sometimes an end is neither an outcome nor a limit, but an activity. Speaking a language and playing a game or musical instrument are ends of this kind. They are not produced by actions, but realized in them. When my end is speaking German, my actions do not, or at least not simply, produce ‘speaking German’; they constitute or realize it. Humanity is also an end of this kind, a *positive* end, that is, something to be realized in various activities. Realization in this sense is making actual what is potential, so we can think of humanity being our (rational) end as making actual whatever potentialities humanity is composed of. And one such potentiality is the second-order disposition to acquire a commitment to the Categorical Imperative.

To summarize, then, our humanity and the disposition of which the good will consists can be thought of as objects of rational (in a Kantian sense) choice in the sense that they are in several, though not all, senses rational ends (which is of course compatible with humanity

also being a means of which we choose to make use). Our humanity – especially in the sense of its including a second-order disposition to acquire the commitment to the Categorical Imperative – and good will can be something to realize in our activities and can be limits on our other ends and activities. If this is so, then we can quite easily make sense of the idea that they are good on the official Kantian theory of value in general. Necessarily for every agent who is rational in the full-blooded Kantian sense, and good will and the humanity in herself and in others are objects of her will in the sense that they are rationally necessary limits on every other end she pursues in every circumstance, as well as are something to be realized and furthered in her actions. They are thus valuable because they are objects of a rational will—because they are related in the right way to a rational will. It is not that their value stems from their being related in the right way to rational choice that distinguishes things of ordinary value from the special value of the good will and humanity. It is the nature of the relation in each case. Humanity and the good will are necessarily and universally the objects of a choice rational in the Kantian sense. Other things are only contingently so. Their value thus need not be seen as a kind of metaphysical glitter.

If I am right that humanity contains the idea of good will (at least understood as the disposition to acquire a commitment any minimally rational will has to the moral law), then we now understand how the good will can be under every circumstance an object of rational choice, and is the condition of which anything else is an object of choice. It is a kind of limit on every other object of choice. Put simply, to say a good will is unconditionally good is to say that I rationally may not choose in a way that will degrade or thwart the development of a commitment to the Categorical Imperative, no matter what other ends I pursue.

#### REFERENCES

- Allison, Henry, 'Morality and Freedom: Kant's Reciprocity Thesis', *The Philosophical Review*, 95, (1986), 393-425.
- Guyer, Paul, 'The Value of Reason and the Value of Freedom', *Ethics*, 109, (1998), 22-35.
- Herman, Barbara, *The Practice of Moral Judgment*, (Cambridge, MA: Harvard U. Press, 1993).
- Hussain, Nadeem, and Shah, Nishi, 'Misunderstanding Metaethics: Korsgaard's Rejection of Realism' in *Oxford Studies in Metaethics* Vol. 1 (New York: Oxford University Press, forthcoming.)
- Johnson, Robert, 'Happiness as a Natural End' in *Kant's Metaphysics of Morals*, M. Timmons, ed., (Oxford: Oxford University Press, 2002), 317-330.
- Kant, Immanuel, *Practical Philosophy*, M. Gregor, tr. (New York: Cambridge University Press, 1996). All page numbers refer to the numbering in the edition of Kant's collected writings, edited by the Koniglich Preussischen Akademie der Wissenschaften, (Berlin, Reimer 1900-), Bd. 1-9.
- Korsgaard, Christine, *Creating the Kingdom of Ends*, (New York: Cambridge University Press, 1998)
- , *The Sources of Normativity*, (New York: Cambridge University Press, 1996)
- , 'The Normativity of Instrumental Reason', in *Ethics and Practical Reason*, ed. G. Cullity and B. Gaut (New York: Oxford University Press, 1997), 215-254.
- Nagel, Thomas, *The View from Nowhere*, (New York: Oxford University Press, 1986).
- Reath, Andrews, 'Legislating the Moral Law', *Nous*, 28 (1994), 436-464.
- Wood, Allen, *Kant's Ethical Thought*, (New York: Cambridge University Press, 1999).

