

**Robert N. Johnson**  
**Was Kant a Virtue Ethicist?**

You might think a simple “No” would suffice as an answer. But there are features of Kant’s ethics that appear to be strikingly similar to virtue oriented views, so striking that some Kantians themselves have argued that Kant’s ethics in fact shares these features with virtue ethics. In what follows, I will argue against this view, though along the way I will acknowledge the features of Kant’s view that make it appear more like a kind of virtue ethics than it really is.

My plan is to first set out the distinctive features of what is nowadays called “virtue ethics”, those features that make it a genuine alternative to other normative theories. I then consider the features Kant’s view might share in common with virtue ethics and the case for saying that it is, therefore, fundamentally the same sort of theory. I follow these two sections with an argument against this position. I want to warn you at the outset, however, that my argument itself will be quite unsurprising, since it is an argument that has been central to the way in which most philosophers have understood Kant’s ethics. Any novelty I can claim here is in my account of what makes virtue ethics a genuine alternative to other normative theories, and my defense of this argument against those, in particular Barbara Herman, who have apparently found the argument unpersuasive.

**I. The Form of Virtue Ethics**

What makes a theory count as a kind of “virtue ethics”? Clearly it is not that the theory takes moral virtue seriously, or that it takes the virtues individually to be an important topic for ethical theory. That would make virtue ethics a trivial addition to an already theoretically sound group of alternatives in ethical theory. Virtue ethicists instead say they offer a genuinely new theoretical alternative, one fundamentally different from deontological or teleological ethical theories.<sup>1</sup> Now one might assume that what makes it fundamentally different is that the theory orients itself around the moral virtue, rather than duty or valuable states of affairs.<sup>2</sup> Thus, instead of viewing virtue as merely an optimistic disposition or disposition to right conduct – that is, as of interest only after the right and the good are understood – virtue ethics gives pride of place to the virtues and these in turn are to inform

---

<sup>1</sup> E.g., see John McDowell, 1997, “Virtue and Reason”, in: Roger Crisp / Michael Slote (eds.), *Virtue Ethics*, Oxford: Oxford University Press, pp. 141 - 162.

<sup>2</sup> E.g., “Certainly it is characteristic of modern virtue ethics that it puts primary emphasis on aretaic or virtue-centered concepts rather than deontic or obligation-centered concepts.” Roger Crisp / Michael Slote, 1997 introduction to *ibid.*, in: Crisp / Slote, 1997, pp. 1-25, p. 3 .

our understanding of the right and the good. That, at any rate, is how some philosophers have regarded virtue ethics.

But this assumption about what makes virtue ethics distinctive is misleading in a number of ways. First, it is not really virtue, but happiness or human flourishing, which has pride of place in these theories, especially those that claim ancient views such as Aristotle's as ancestors. Indeed, flourishing has to be foundational because it is the source of the list of virtues: The virtues are just those character traits that are integral to a flourishing life. Second, although human flourishing has pride of place in virtue ethics, it also is not really, or at least not precisely, what makes virtue ethics fundamentally different from the others. After all, utilitarianism also puts human flourishing at the center of its theory. What makes virtue ethics different from utilitarianism is in its understanding of the nature and role of human flourishing in a normative theory. For utilitarianism, flourishing is a valuable state to be furthered by our actions. For virtue ethics, flourishing is an activity, that of "doing well" as a human being. The latter conception of flourishing as an activity shows that virtue ethics has a radically different understanding of its role in normative theory. Hence, it is really the role of human flourishing plays in the theory that makes virtue ethics distinctive.

More precisely, what makes theories collected under the heading of "virtue ethics" fundamentally different theoretical alternatives, if they are such, is, first and foremost, the form these theories take. What I mean by the "form" of a theory is how it arranges and argues for positions regarding the main topics of ethical theory. Those topics are the nature of right and wrong, of good and bad, and of what makes for a life worth living. Of course some philosophers, call them "pluralists", would argue that there are no systematic interconnections between the topics. But most hold that some one of these topics is the primary and the other two secondary or derivative. While most ethical theories approaches these topics by treating the good or the right as foundational, the alternative to these would build its ethics on what makes for a well-lived life. Hence, the virtues would only enter the picture theoretically once a conception is in place of what makes for such a life, or as I shall sometimes call it, an ideal of the person.

The claim that an ideal of the person, rather than duty or value, is the foundation of ethics is not, or at least not mainly, an epistemic claim. The central claim is not that necessarily one cannot know or have justified beliefs about which actions are right or which states of affairs are good without first knowing or having justified beliefs about what constitutes such an ideal. It could well be that we cannot know what a well-lived life is without first understanding moral duty and what is genuinely of value. But it may still be the

case that what makes an action a moral duty and what makes something of value is its relationship to some ideal of the person. That may just be the only epistemic route to that grounding ideal. Nor is the claim necessarily conceptual or semantic, that is, a denial that concepts of this ideal are definable in deontic or value terms, or that statements about the ideal are true in virtue of obligation or value facts. Some might think that you can only define a well-lived life in terms of doing one's duty, others might well claim it goes the other way around, but that is not the same thing as arguing about which is foundational to theory. Moreover, the denial that right and value are foundational is also not primarily a *metaphysical* position, for instance, that an ideal of the person is "more real" or more capable of independent existence than the properties of right and wrong action or good and bad.

Virtue ethics is primarily a *normative* view and hence the central position has to do with what is basic *in a normative sense*. This means that the answer to traditional normative questions about the difference between right and wrong, which states of affairs are good or bad, and what makes lives worthwhile, are, on the virtue ethical approach, to be arranged such that the last issue forms a basis for answers to the first two. Those actions are right and those things good that are related in the right ways to an ideal of the person, or so virtue ethics should say. Hence, you should act in the ways that characterizes that ideal, and valuable states of affairs are just those that characteristically come from living that ideal. To be sure, pluralists will reject the idea that any of these notions should be the basis of the others. But those virtue ethicists who think there should be some account of right action and of good states of affairs typically have in mind accounts that appeal to flourishing, as an ideal of how to live, as the foundation for right action and valuable states.

These features of the form of virtue ethics in fact show it to be, not *the* novel alternative to deontological or teleological ethics, but *one possible instance* of such an alternative. For were one to use the central claim that human flourishing is foundational to ethical theory to generate some account of right action, it would look something like this:

VE: For all actions  $\phi$  and all persons S, it is right (to be done, ethical, correct, etc.) for S to  $\phi$  in C at  $t$  if and only if  $\phi$ ing in C at  $t$  is or would be characteristic of a flourishing human life.<sup>3</sup>

Yet VE is an instance of a more general schema of this form:

V: For all  $\phi$  and all S, it is right (to be done, ethical, correct, etc.) for S to  $\phi$  if and only if  $\phi$ ing is (or would be) characteristic of P.

---

<sup>3</sup> Hereafter, I omit the spatiotemporal markers for clarity's sake. The formulation is roughly what one finds in, for instance, Rosalind Hursthouse (see Rosalind Hursthouse, 1999, *On Virtue Ethics*, Oxford: Oxford University Press, p. 28); McDowell, 1997; and of course in Aristotle, 1993, *Nicomachean Ethics*, trans. by T. R. Irwin, Indianapolis: Hackett, 1105b6.

where “P” stands for some ideal of the person. To assume that VE represents the alternative to deontological or teleological approaches to ethics is to assume too much already, that the only ideal possible is one that would fit under the heading of *eudaimonia*. But it is, or at least can be, a substantive issue what sort of life is ideal and whether that ideal is a *eudaimonic* life. Hence, strictly speaking it is V, not VE, that is the alternative to deontological and teleological approaches to ethics. It is a bi-conditional according to which the left-hand side speaks of actions in deontic terms such as “right”, and on the right-hand side speaks of the behavior characteristic of some ideal. The alternative theory says that we should generate an account of right action by appeal to the conduct characteristic of that ideal. For Aristotelians, P is the *eudaimonic* or flourishing person. But others, armed with a different ideal, would come up with a different set of actions that are right or to be done on V. What all such views share is the notion that an ideal, however that idea is spelled out, is the standard of correct conduct (and valuable outcomes), however that idea is spelled out. Such ideals (and note well that I leave it open *how* ideal they should be) are undoubtedly a central subject of ethical theory. But it would be a mistake to assume that all such ideals must have something to do with the virtues or more generally with human flourishing, as normally understood in discussions of virtue ethics.

The form of V, then, is a critical component of virtue ethics, inasmuch as it is supposed to represent an alternative to deontological and teleological theories. But this is not yet sufficient to make it a genuine alternative. Utilitarianism and intuitionism, for instance, both could accept V, plug in their favored accounts of “right action” on the left hand side, and generate on the right hand side an ideal characterized by that sort of conduct. That, then, would be the theory’s ideal of the person. The notion of such an ideal would simply be motivated by a prior notion of right action. So it is essential, if it is to be a genuine alternative, that virtue ethics makes the right hand side as far as possible *independent* of the left. That is, if virtue ethics is to be distinctive, the argument for the particular view of an ideal of the person (much less a view of human flourishing or happiness) must be innocent of premises presupposing some particular understanding of right action or good states of affairs. If it is not, then it would be simply building into the right hand side the elements of right action of which the ideal is supposed to be the source. The ideal of the person in the theory will be doing no work in generating answers about how to act.

The sense in which I mean that the ideal of the person “generates answers” is that it gives a rationale for particular lines of conduct; this will become clearer as we proceed. For now, notice that the other orientation of V, which takes the right as basic, “generates answers”

about the ideal of the person at least in this sense: The question “What sort of ideal should I live up to?” gets answered constrained by first answering questions such as “What must I do?” and “What must I not do?” Whatever ideal conforms to the answers to those questions alone is ideal for those who take the left hand side as normatively basic.

## II. The case for Kantian virtue ethics

The account of the unique theoretical orientation of virtue ethics I’ve just offered consists of two features. First, it ties right action to an ideal of the person in the way that the bi-conditional in V does. Second, it generates the account of right action on the basis of an account of that ideal as far as possible by understanding the latter account *independently* of the notion of right action (or good states of affairs). Assuming this account is correct, surprising as it may seem, one might naturally view Kant’s own ethical theory as sharing both of these features with virtue ethics. Here is why.

The central idea is that the notion of a “good will” that inaugurates the *Grundlegung* and in particular the argument of *Grundlegung I*, offers just the form that supports this way of understanding Kant’s ethics. Kant begins by specifying the nature of the good will apparently as if it characterizes an ideal of the person for rational agents such as ourselves, an ideal whose value is beyond measure and which is to be honored and respected. He then appears to use this as a basis on which to generate a rule of dutiful action, the Categorical Imperative. The idea, at least as it appears to be, is that it is characteristic of a good will to conform to the Categorical Imperative, and so this must be the fundamental principle of right or the nature of duty on the Kantian account. Thus, the assumption is that dutiful acts are those characteristic of a good will. And it is this assumption that has the form of V:

KE: For all  $\phi$  and S, it is S’s duty to  $\phi$  if and only  $\phi$ ing is (or would be) conduct characteristic of a good will.

KE thus paints Kant’s ethics as having precisely the same form as virtue ethics, the form of V, in which P is Kant’s good will. Bringing to bear features of the argument of *Grundlegung I*, we are now also in a position to build a case for the second aspect of a virtue ethical theory, namely, that the disposition of a good will is the foundation of dutiful conduct in the way that the ideal is the foundation of right conduct in V. For the structure of that argument appears to support the view that Kant’s ethics is a form of virtue ethics. Whether Kant’s view is a kind of virtue ethical view will turn on the second crucial component of those views, the component that grounds the right in the ideal on the right hand side of V. Hence, it turns on whether the good will in Kant’s ethics is in the right sense a foundation for the Categorical Imperative.

In what way might the good will – as an ideal on the model of ideals such as the Aristotelian *phronimos* – provide a foundation for the distinction between right and wrong conduct? If it is to play the foundational role the ideal does in virtue ethics, it must play the role of a final end. A final end in such theories is supposed to provide a grounding rationale for our behavior, a guide to deliberation that lays down a rule, principle or law for our behavior. The rationale, then, for Kant's ethics would be that the CI is laid down as a law for us because the value of the good will as a final end, an ideal of rational nature let us say, demands this. The feature of the conduct violating the CI and making this constraint understandable is that the conduct is uncharacteristic of a good will.

Ends could be the source of practical laws in the following way: For Kant, a person's end is the reason she wills an action; as such, it is the ground of the laws guiding the will of that person. If it is not an end in itself, but an end only for some rational agents and not others, the law guiding the person's deliberation is only a hypothetical imperative, to take the means to that end or else abandon the end. Something that is an end in itself, by contrast, would be an end for everyone. Hence, it would have to be a reason for a different sort of law guiding deliberation, a universally valid categorical imperative.<sup>4</sup> If the good will, and more generally, the humanity in persons, is an end in itself, then it seems we have what we are looking for: Kant's ethics grounds the authority of the Categorical Imperative in a final end, the good will and (or) the humanity in persons (understood as including the capacity for a good will). The gulf between virtue ethicists and Kantians suddenly appears to be not nearly as deep as many may have assumed. The issue between the two views concerns only the content of an ideal for human beings, not whether some such ideal or other should be theoretically fundamental.

Before I assess this line of thought, there is one feature of KE I think that it is worth noting is a potential advantage for Kantian ethics over other doctrines that accept the form of V. The good will, as Kant conceives of this, is not necessarily a *virtuous will*. The lack of moral virtue, which Kant holds is the lack of moral strength of will to overcome obstacles to doing one's duty, is compatible with nevertheless having a good will.<sup>5</sup> Because of this, the ideal P as it appears in Kantian ethics is compatible with non-ideal character traits. That in turn allows us to ask a critical question for moral theory, namely, What would a good will characteristically do to improve his own character? The answer will then deliver, on the left hand side of the Kantian version of V, a duty to improve our own characters. I say this

---

<sup>4</sup> G 4:427-8. All references to Kant's works are to Immanuel Kant, *Practical Philosophy*, ed. and trans. by M. J. Gregor, 1996, Cambridge: Cambridge University Press, 1996

<sup>5</sup>“Weakness in the use of one's understanding coupled with strength of one's emotions is only a *lack of virtue* and, as it were, something childish and weak, which can indeed coexist with the best will.” MM 6:408

because, surprising as it may seem, any view that characterizes the ideal of the person as an ideal of *character* – as someone who by contrast *is* in possession of all of the virtues – will not be able to explain why we have a duty to improve our characters. It will not be able to explain this because someone whose character is already completely virtuous – the very ideal on the right hand side of V in terms of which the theory is trying to generate a set of duties – would not characteristically engage in character-improving behavior.<sup>6</sup> This fact about Kant’s theory, that his ideal of the good will is explicitly *not* cast in terms of virtues is, to be sure, an important difference between Kant’s ethics and virtue ethics. But, as I have argued above, the role of the virtues in virtue ethics is not the source of its unique theoretical orientation. Indeed, if I am right, the dominant role of the virtues in such theories may even be a serious liability. As long as the ideal motivating the theory of right is conceived of as an ideal of character, the theory will not be able to explain duties of self-development.

I realize that in the eyes of many this fact alone will disqualify Kant’s view as counting as a kind of virtue ethics; after all, virtue – either as a generalized disposition or as a set of particular traits – is not at all central to his theory. But if I am right, virtue is not really so central to theories called “virtue ethics” either. The critical move in these views is to draw our attention to the idea of the person, to its centrality and importance to our ethical concerns, and to challenge the notion that value or duty could occupy a place of that prominence in practical deliberation. That move does, to be sure, lead very naturally to a consideration of the enduring traits of character that are necessary to realize such a life. After all, so much of our own concerns with children are with instilling the sorts of habits and dispositions that will ensure that they have an opportunity to live a worthwhile life. Nonetheless, these considerations follow upon an original focus on the ideal of the person, an ideal about which Kant can be said to have an in-house disagreement with virtue ethicists.

### **III. Leaving Virtue Ethics Behind**

My discussion so far has laid out the two features of virtue ethical theories that mark it out as a genuine alternative to other theories, and has offered what I take to be the reasons for thinking Kant’s view shares those features. In this section, I explain why, nevertheless, Kant’s position is wholly incompatible with a critical feature of virtue ethics, why, that is, Kant’s orientation is as deontological as we always thought it was.

KE is itself not, so far as I can see, incompatible with Kant’s views. So let us take as read that our duty is to conform to whatever principle that a good will embraces. So to this

---

<sup>6</sup> I argue this more fully in Robert Johnson, 2003, "Virtue and Right", in: *Ethics* vol. 113, pp. 810-834.

extent, we should acknowledge that Kant's ethics has the same form as virtue ethics. But this is only half the story. The critical question is whether the second element, the element that makes the right hand side of theories of the form of V basic, is also a part of Kant's views. I'll argue that this second element is quite incompatible with his ethics.

First, if the right hand side of KE is basic, then Kant must be able to characterize a good will independently of the principle of duty it supposedly generates. That is part of the idea of making the right hand side basic. But clearly this is wrong. There is no way of characterizing what a good will is, in Kant's moral philosophy, without deploying the idea of the Categorical Imperative. To be sure, appearances are deceiving: The order of presentation of ideas in the *Grundlegung* appears to set out the good will as something we already fully and clearly grasp. But in fact, this is only a rhetorical device, a device through which Kant can display how it is that we discover the fundamental principle of morality. Ordinary moral consciousness has a stark intuition of what it is that makes for a person of good will in his view, but even Kant would admit that that consciousness would not represent the Categorical Imperative so evidently as it's guide. Nevertheless, Kant is confident that this is the principle each person deploys when he considers what he must do.<sup>7</sup>

Second, this order of discovery is most definitely not an order of normative priority. Kant is, after all, engaged a *search for* the supreme principle of morality that Kant is undertaking in the first two sections of the *Grundlegung*.<sup>8</sup> In order to *discover* the fundamental principle of duty, Kant assumes that this principle will be one and the same as the principle of a good will. That is, he assumes something like KE. But this assumption is based on the idea that what makes a good will good is her disposition to deliberate and decide what to do by way of the fundamental principle of morality (whatever that principle turns out to be). The authority of the CI, the property in virtue of which it inspires respect in us, is not the fact that it is the principle of a good will. Quite the reverse. It is because of her embrace of the authority of the CI that her will enjoys the special value it possesses. Thus, whether or not the good will represents an ideal of some sort, it is not a fundamental notion out of which a theory of right action can be generated.

Third, the passages in which Kant argues that an end in itself – the humanity in persons – is a ground of the Categorical Imperative, as well as his discussion of the value of the good will, must one and all be understood against the backdrop of his general view of value. For instance, take perhaps the most well known passage that sets out the importance of the value of humanity:

---

<sup>7</sup> E.g., see G 4:403

<sup>8</sup> G 4:392

“But suppose there were something the *existence of which in itself* has an absolute worth, something which as *an end in itself* could be a ground of determinate laws; then in it, and in it alone, would lie the ground of a possible categorical imperative, that is, of a practical law... Now I say that the human being and in general every rational being *exists* as an end in itself...”<sup>9</sup>

To be sure, this passage appears to state that there is a unique value that is the basis of the moral law. Yet, as strong and clear as these words seem, they do not, in fact, say this. To have value, in Kant’s view, is to be the object of rational willing.<sup>10</sup> Hence, for something to have absolute worth is for it to be, as it were, absolutely – or in every possible circumstance – the object of rational willing. So although the discovery of something with absolute worth is proof that there is a categorical imperative – an imperative that necessarily binds all rational agents – this does not show that the direction of normative priority in KE is from right to left. It does not show, that is, that it is the value of rational nature, the value of the good will, or any other value, that is the source of the categorical imperative. What it shows is that there is something that must be the object of any rational agent’s will, and as such absolutely good. And if there is something that must be the object of any rational agent’s will, whatever determines that will to make that thing its object would be a categorical imperative. That our wills are bound by a categorical imperative is why humanity and the good will have the value that they do, not the other way around.

Another way to put the point is this: It is not simply the *existence* of an end that lays down a law for deliberation and decision; it is the *willing* of an end that lays down a law.<sup>11</sup> Being an end is nothing over and above being the object of rational willing. And so being an end in itself is nothing over and above being *in itself* an object of rational willing. The reason *willing* an end give one a law is because rational willing is, in Kant’s view, autonomous. A rational will must, by virtue of being a rational will, be committed to universal law. To be sure, Kant’s position is astonishing and controversial. But it is at the very heart of his ethical thought. To think that there must be some value to be a ground for conforming one’s will to moral law is to think that a rational will is heteronomous.

These considerations against a virtue ethical reading of KE are not *recherché*. They are elements of a standard understanding of Kant’s ethics. Yet some Kantians have argued that, nevertheless, this standard understanding is entirely wrong, that Kant’s ethics shares the unique theoretical orientation of virtue ethics I set out above. In particular, Barbara Herman has argued that Kant’s view is not a “traditional deontology” in which there is an “absolute

---

<sup>9</sup> G 4:428

<sup>10</sup> CPrR 5:58-65

<sup>11</sup> MM 6:385

priority of the right”.<sup>12</sup> Deontology, she argues, has no conception of value at its foundation.

And that means that there is nothing in the theory to offer a “rationale” that

“renders moral action intelligible to the agent making possible the reasoned integration of morality into one’s system of ends [and] introduces a framework for reasoned deliberation necessary to the stable resolution of morally complex situations. A grounding conception of value could provide this rationale by offering an explanation of the wrong- or right-making characteristics of action that renders moral requirements intelligible in a way that is then able to guide deliberation...One way of putting the question to Kantian ethics is to ask whether the unconditioned good—the good will—can play the didactic role of a (the) final end.”<sup>13</sup>

Herman understands the priority of the good as its playing a “didactic” role of a final end, an end that can guide deliberation and allow moral constraint to make sense to the agent in terms of her own system of ends. She is clearly not speaking of “the good” as a state of affairs to be produced by us, and so is not suggesting that Kant’s project is parallel to consequentialism. She is not, in particular, supposing that conforming to the CI is required of us because in so doing we will produce a good will in ourselves. She is thinking of the good will as occupying the selfsame position in a theory that virtue ethicists think of *eudaimonia* or human flourishing as holding. In her view “rational nature is...the regulative and unconditioned *end* of willing—that is, a final end, an end-in-itself”.<sup>14</sup> What makes this possible is that rational nature is “autonomous, having the capacity to be its own original source of reasons.” This makes rational nature “a unique kind of value”, the kind of value that can play the didactic roles Herman, and she supposes, Kant, sets out for it.<sup>15</sup> The basic idea is that you need to be able to say why it is good to conform to some rule if that rule is to guide your deliberations and decisions. But, according to deontological views, you are not able to say this with regard to moral rules. Luckily, in her view, the whole account of duty in Kant’s ethics is in fact based on a conception of value, and that is the conception of the good will (and the related notion of humanity in persons).

Herman does not cast her view in as clear a language as one might like. She says that “we can understand the formal requirements of practical reason as a conception of value”<sup>16</sup> and that “we cannot understand what practical reason is without understanding it as a conception of value”<sup>17</sup>; the “purely formal principles of rationality can be expressions

---

<sup>12</sup> Barbara Herman, 1993, “Leaving Deontology Behind”, in: Barbara Herman, *The Practice of Moral Judgment*, Cambridge MA: Harvard University Press, pp. 208 – 240, p. 210.

<sup>13</sup> *ibid.* p. 216.

<sup>14</sup> *ibid.*

<sup>15</sup> *ibid.* p. 238.

<sup>16</sup> *ibid.* p. 239.

<sup>17</sup> *ibid.*

of...conceptions of value".<sup>18</sup> But neither a principle of rationality nor the faculty of practical reason is a conception or expression of anything. We *could* understand her point as being that a principle, in binding our will, does so by stating what it would be good to will, or else that rational principles are standards of valuable behavior, standards of what it would be good to do. But if this were her meaning, then there would be no reason to suppose that this is not a deontology. It is just to restate Kant's view that the good is the object of a rationally determined will, a will determined by principles of practical reason. Those principles, or at least the categorical imperative, cannot themselves be, in turn, based on some value, since embracing those principles is supposed to compose what willing anything at all consists in.

In fact, Herman anyway explicitly rejects this way of understanding her point.<sup>19</sup> She wants to know why conforming to the principles of rationality is good, how rationality itself is a value.<sup>20</sup> I can only think that this must just be the question, What reason do we have to be rational? Hence, if Kant's view holds that we must conform to the Categorical Imperative because it is a demand of rationality, we now need to know what reason there is to be rational – what good is it to be rational. Here, however, we have to distinguish between the faculty of rationality and the various principles that characterize rational willing. It may be that, as for instance John Broome has argued, there is reason to possess the faculty of rationality – i.e., that set of dispositions to conform to various principles of rationality – but no reason to satisfy any individual principle of rationality.<sup>21</sup> It may be so because while it may be in our interest to be rational, this does not show, and in Broome's view there seems to be no way to show, that there is a reason to satisfying any individual principle. They are two different questions.

I think that although Herman often states her question as what rationale there can be for moral constraints, I believe she is not looking for a reason to conform to the Categorical Imperative – an individual principle of rationality. What Herman wants is a reason to possess the faculty – characterized, to be sure, as Kant uniquely characterizes it – of rationality, since she speaks of the set of rational principles as a conception of the good. She wants to know what good there is in possessing *that* faculty of rationality, characterized as encompassing dispositions to follow moral as well as instrumental principles. But, unlike Broome, she does not want an answer in terms of our interests. In particular, if the rationale for conforming to moral constraints is simply that we must be rational – possess the disposition to conform to such constraints as well as the other dispositions – then we need some way to understand

---

<sup>18</sup> *ibid.* p. 236.

<sup>19</sup> *ibid.* 215.

<sup>20</sup> *ibid.* p. 213.

<sup>21</sup> John Broome, 2005, "Does rationality give us reasons?", in: *Philosophical Issues*, vol. 15, pp. 321–37.

being rational as intrinsically valuable. It must, as *eudaimonia* does in the Aristotelian theory, all by itself end questions about the point or purpose of action. Unless we have this endpoint of deliberation, she thinks, we are left unable to resolve deliberative conflicts, such as when morality asks us to act in ways that are deeply troubling to us. What rationale is there for a moral constraint that asks us to give up things of such great importance that to do so will seemingly damage ourselves or those we care for? We must be able to see it as in some way the realization of a certain kind of intrinsically valuable identity as a rational agent.

The crucial move in this story is to appeal to rationality as a kind of intrinsic value. But Herman's ending the account at this point cannot be squared with Kant's overall theory of value. Although she points out that things come to have value in Kant's view by being the objects of rational willing, she apparently does not regard this as true also of things that are intrinsically and unconditionally good. But this is a mistake. To be intrinsically and unconditionally good is also to be the object of rational willing. There can be no *independently existing* value in Kant's value theory, not even humanity and the good will.<sup>22</sup> Humanity and the good will are ends in the sense that they are to be respected in the ways Kant discusses. But it is not that Kant is pointing out that these are a kind of non-natural gem existing in the world whose value demands certain attitudes and behavior on our part. To have value is nothing over and above being the object of these sorts of rational demands.

There is an important reason why Kant's theory cannot be read in the way Herman wants to read it. As she points out, of course, that theory denies that there is a *nonmoral* rationale for moral constraint. Hence, it denies that moral action needs "the reasoned integration of morality into one's system of ends", *if* those ends are understood as *nonmoral* ends. That is not the failure Herman is concerned with. But the whole point of a deontological view is to argue that moral agents *already perceive or embrace* moral principles. It thus follows that if you are rational enough to know you have a duty, you already will to conform to it. That is why there is nothing in Kant's ethical philosophy to answer the question "Why be moral?" Any minimally rational agent in his view *already* possesses a will that, by its very constitution, wills the moral law. That distinguishes it from deontological views that claim that any rational agent must already *perceive* the moral order.

It is thus central to Kant's position that moral philosophy *not* address questions about the rationale for moral constraint, even if it must address why, if there are moral constraints at all, every rational agent is already in touch with them. This is what makes sense of Kant's position that, unlike those such as Hobbes and Hume who came before him, there is no need

---

<sup>22</sup> I argue this point in some detail in: "Value and Autonomy in Kantian Ethics", forthcoming in: *Oxford Studies in Metaethics*, vol. 2.

to set out a reason to conform to the Categorical Imperative. It is not the fact that those theories try to tell us why, in terms of our *subjective contingent* ends, we should conform to moral restrictions that he rejected them and regarded them as “heteronomous” theories.<sup>23</sup> What makes them heteronomous is that their very conception of the rational agent are of an agent whose will does not already commit himself to the moral law. If that is right, then wanting to give a rationale for moral constraint to rational agents makes less sense than giving Yankees fans a rationale for wanting the Yankees to win.

It might now be thought that if, as I say, a central feature of Kant’s ethics is a conception of persons as already committed to the moral law, then his theory, after all, *does* ground itself in an ideal of the person. It is an ideal of the rational agent as already committed to the moral law. But, first, the conception of rational agency here does not play the role in the theory that it would have to play in order to provide a rationale for moral constraint. The role Kant’s conception of agency plays, instead, as I explain above, makes such a rationale otiose. Second, Kant’s conception of rational agents is itself wholly informed by a prior understanding of the moral law. As such, it obviously cannot be then used to provide a rationale for the moral law.

#### **IV. Conclusion**

I said that the beginning that my point would not be surprising. Although there are theories that provide a final end to play the didactic role in deliberation that Herman sets out, I think Herman is quite wrong in arguing that Kant’s could be one of them. Kant does not offer anything to make morality intelligible to someone whose ends are not already arranged around, and whose deliberations are not already informed by, the moral law.

---

<sup>23</sup> 4:441-445