

Kant's Formula of Humanity

Gary Banham, Editor, *Kant Studies Online*

There has been an increasing tendency in contemporary work on Kantian ethics to focus on the formula of humanity, partly because it is thought that his formulas of universal law are insufficient to underpin demanding ethical requirements and also because it is taken to be the case that the Formula of Humanity states a substantive value that has wide appeal and resonance. Two basic notions are derived from the humanity formula, one, that we should treat persons with respect, and, two, that we should not treat them as means except insofar as we also view them as ends-in-themselves. These two requirements can be termed the “respect requirement” and the “mere means” requirement. However, whilst there is considerable agreement concerning the importance of both these requirements there are problems concerning both how to interpret their application and how to understand the meaning both of these requirements and of the Formula of Humanity as a whole. I will focus here on the second question concerning the meaning of the Formula of Humanity and its relation to these requirements, since, if we do not get this right, it is unlikely we will arrive at a correct view of how to apply either the Formula or these requirements.

The Capacity of Humanity

The Formula of Humanity refers to using humanity “in your person” and prior to introducing the formula Kant has made manifest that what he is referring to by the notion of “humanity” is not something species-specific but rather a property that would belong to all rational beings in general. The first question, in assessing the meaning of the formula, is thus to ask what capacity of

rational beings Kant is picking out when he refers in this formula to “humanity”. Since “humanity” is a property of rational beings in general it would make sense to think of this capacity simply as that of rational nature itself though this leaves Kant open to the objection, as formulated by Derek Parfit, that this would mean that “it would be wrong for us to damage our ability to play chess or solve crossword puzzles” since these are activities that manifest rationality in a generic sense.¹ So, Kant must naturally mean, if his reference to “humanity” involves a claim about rational beings in general, more than a generic capacity for rationality. Allen Wood comes close to stating a view like this minimal one that Parfit objects to when he claims that “the sole fundamental and unconditional value is the value of rational nature as an end in itself” though he buttresses it when he adds that our basic act as rational beings is “the act of setting ends and regarding them as good”.²

Wood’s description of the capacity Kant is terming “humanity” runs together two distinct ways of describing it. The conception that the capacity Kant terms “humanity” is simply the capacity of “setting ends” has been embraced by Christine Korsgaard and seems to imply that it is rationality considered as the ability to be purposive and have instrumental aims that Kant is viewing as an “end-in-itself”. However, Korsgaard’s conception becomes more complicated when we note that she takes Kant to have the view that ends play two distinct roles in the determination of conduct, either as purposes to be pursued, or as something one should not act against.³ Now, if there are ends that one should not act against, it follows that, despite Korsgaard’s “official” description of the capacity of humanity, as simply the ability

¹ Derek Parfit (2011) *On What Matters* (Oxford University Press: Oxford and New York), Volume 1, p. 242.

² Allen Wood (2008) *Kantian Ethics* (Cambridge University Press: Cambridge and New York), p. 92.

³ Christine Korsgaard (1986) “Kant’s Formula of Humanity” in C. Korsgaard (1996) *Creating the Kingdom of Ends* (Cambridge University Press: Cambridge and New York), where she gives the generic account of purposiveness on p. 110 but the two senses of “end” precede this on p. 108.

to set purposive instrumental aims, that this claim does not reflect her investigation of how purposes work in the discussion of Kant's formula of humanity.

Noting this discrepancy between Korsgaard's generic description of the capacity and her more specific distinction between types of "ends" helps to bring out the move that happens when Wood incorporates into his description of "humanity" not merely the capacity of setting ends but also the reflexive ability of being able to describe them as good. If there is something involved in this ability that is an important part of the capacity of "humanity", however, it follows that this capacity is not equivalent merely to purposiveness in general, but must also include some restrictions on the exercise of purposiveness. Korsgaard confesses as much when she amends her view in recognition of Kant's claims about the good will being of unconditional value, a point that would conflict with her more minimal "official" claim about purposiveness in general. In recognition of the status of the good will Korsgaard states that whilst humanity as a general capacity is equivalent to the power of rational choice that it is "only when the choice is fully rational is humanity fully realized".⁴ But this concession involves including at least one and potentially two new elements into her account over and above the "official" one of reference to purposiveness in general. One element would be a *normative* sense of rationality sufficient to respond to the obvious objection that I mentioned earlier from Parfit and that would, in building in a sense of rationality that was not purely descriptive, have to indicate something that was of value in purposiveness over and above the mere act of setting ends. Secondly, it is also possible to view Korsgaard's reference to "full" rationality being "fully" realized as containing a notion of maximisation. I am not claiming she intends the second implication, though, as we shall see subsequently, others certainly do intend it when they make a point similar to Korsgaard's.

⁴ Korsgaard, *op.cit.*, p. 123.

If, however, we assume, following Korsgaard's second suggestion, that there is something to the idea of "humanity" being more or less fully realized, then we will need a description of the way in which this can take place that somehow reflects her point about the greater or lesser normative rationality present in choices. This brings us back to Korsgaard's suggestion that there are certain types of ends that we should not act against inasmuch as we possess the property of humanity and which Kant himself stresses when he describes the formula of humanity as stating that an end in itself "must serve for every maxim as the limiting condition of all merely relative and arbitrary ends" (Ak. 4: 436). In understanding Kant's discussion of purposes, as inclusive of this idea of ends that we should not act against, and, which limit other ends, we arrive at a sense of "humanity" as containing in itself a kind of end that is not like other ends and which has some kind of authority over them.

So part of the point of the formula of humanity is to bring out this sense of an end that limits the degree to which all other ends can be adopted or pursued. In this respect the formula of humanity appears to pick out a negative requirement. This leads John Rawls to depict the capacity of humanity as nothing other than "animated pure practical reason" as it is descriptive of "powers and capacities" that include what he terms "moral personality", described as what makes it possible for us to have a good will and a good character. However, like Korsgaard, Rawls runs this account in tandem with the more minimal one of simply seeing humanity as that "in virtue of which we are capable of setting ends".⁵ Like Korsgaard, Rawls appears to be capable of assuming that the mere capacity to set ends is normatively distinctive but, when pressed, views the setting of ends as part of a general moral character and hence departs from the

⁵ John Rawls (2000) *Lectures on the History of Moral Philosophy* (ed. by B. Herman, Harvard University Press: Cambridge, Mass and London), pp. 188-9 where these differing views are stated within lines of each other without Rawls appearing to notice any tension between them.

depiction of ends in terms that would collapse them into being considered merely as part of a description of instrumental reason. If Rawls' fuller account is considered, however, the reference to pure practical reason and moral personality suggests a close connection between humanity and the good will.

Two other views of the capacity for humanity help us to bring together some of the characterisations we have seen thus far. Richard Dean assumes that the property of "humanity" is normatively loaded and he refers to it as an "ideal" toward which we should strive, hence, not even as a settled property that is possessed by us. On Dean's account what makes "humanity" descriptive of a property we can term an "end in itself" is that it is rational nature that *includes* "some aspect of moral reason – either the power to legislate moral principles, the overall capacity to act on moral principles, or the commitment to actually act on them".⁶ Now, of Dean's three candidates, there are clear problems with the second two. If "humanity" is meant as either the overall capacity to act on moral principles or the commitment to actually act on them then it appears that the nature of such principles is something that should really merit the name of "end in itself" since it is not obvious why these capacities or commitments should be taken as what is ultimately valuable. That would leave the power to legislate as the most likely candidate for the capacity of humanity being taken as a normatively significant form of rationality that could also be more or less progressively realised.

A final candidate for the capacity of humanity that is on offer is that provided by Paul Guyer who regresses on the kind of account I have suggested follows best from Dean's view. This involves tracing the capacity to legislate moral principles as an expression of freedom and thus viewing the property of "humanity" as nothing other than another term for "freedom". Guyer puts this view when he writes:

⁶ Richard Dean (2009) "The Formula of Humanity As An End in Itself" in T.E. Hill (ed.) (2009) *The Blackwell Guide to Kant's Ethics* (Wiley-Blackwell: Oxford), p. 87.

“Recognising our freedom, under the name of our humanity, as an end in itself that has absolute rather than contingent worth gives real point to the formal requirement of acting on maxims that can be universal law; it requires that our maxims can be accepted by all as preserving the freedom of all”.⁷ On Guyer’s view the reference to humanity indicates possession of rational freedom and there are types of maxims that are more or less consistent with the possession of this freedom. If “humanity” is taken as equivalent to the possession of such freedom, and there are maxims that are more or less consistent with this, then this does give us both a means to view its rationality as inclusive of a normative property and a description of what it is that can be more or less fully realised. Both these points are useful elements of Guyer’s account but the conception of freedom as a normative property has the obvious difficulty that it appears to mean that wrong actions are, as such, un-free and thus to create a problem for attributing responsibility for such acts. As I will come back to subsequently, Guyer’s account also openly courts the danger of opening the door to a kind of maximisation procedure that I indicated could follow from Korsgaard’s conception.

So we have a number of possible options for the characterisation of the property of “humanity” and all these have significant problems. If we take this property to be merely the capacity to set ends, it appears too minimal to describe something that we could term an “end in itself” and to collapse into a description of instrumentality. If, however, we add further normative properties in addition to this general structure of purposiveness, such that we arrive at a sense of something which has authority over all other ends, then we need an account of what possesses this structural property of being able to set ends that are not merely contin-

⁷ Paul Guyer (2002) “Ends of Reason and Ends of Nature: The Place of Teleology in Kant’s Ethics” in P. Guyer (2005) *Kant’s System of Nature and Freedom: Selected Essays* (Oxford University Press: Oxford and New York), p. 175. Notably, and doubtless not coincidentally, Guyer’s phraseology here closely relates to that of Kant’s in the Doctrine of Right.

gent but necessary in themselves. If, however, we take Guyer's route and view this property purely as freedom itself then we appear to arrive at a problem with how to view bad actions as actions that agents are accountable for.

Autonomy and Humanity

Having gone through these distinct accounts concerning what the property of "humanity" consists in I want to offer one that combines the elements given and, in doing so, avoids the problems of each of them. The minimal description of the capacity of "humanity" simply as the setting of ends in general is to identify it with the condition that Kant describes as that which enables there to be imperatives at all. The importance of the reference to purposiveness is that it enables Kant to describe not rational nature generally but rather rational willing as he takes rational willing to involve the setting of ends and the adoption of means that are congruent with the achievement of the ends set. However, Kant also distinguishes between "subjective" and "objective" ends, meaning by the former ends we adopt because, for whatever reason, we happen to desire them. Kant also describes these ends as "material" since they are ones that are adopted on the basis of needs and desires. By contrast, "objective" ends are termed by Kant "formal" since they are not concerned with desires or needs. The material, subjective, ends all manifest the structure of purposiveness in general and are capable of imperative form, the form Kant terms "hypothetical" imperatives, taking the general structure that states if you want x, do y as y is the means of attaining x.⁸

The point about the subjective, material ends is that they are contingent as they are conditional on the presence of empirical factors and, not only are they so contingent, they are also fungible, that is, it is possible to trade one of these

⁸ Though, as I will indicate in conclusion, sometimes adopting such a "direct" strategy is not appropriate when following hypothetical imperatives.

ends off in favour of another. In this sense, subjective ends all have what Kant terms a “price” and are capable of being acquired. By contrast, objective ends are necessary ends and they are not capable of being converted according to any common currency, they are thus possessed of “dignity” rather than “price” and a “dignity” that is, indeed, in Kant’s terms, *beyond* all price. This objective, necessary, end is what Kant terms the “humanity” in oneself and it is related by him to the categorical imperative in the formula of humanity. Now, given that it is related to the categorical imperative in this way, it follows that action with regard to “humanity” has some universal requirements and these were specified in Korsgaard’s notion that “humanity”, in some important way, sets a limit on all other ends that can be pursued. The description of universal law that is given both as the supposedly “single” categorical imperative and in the law of nature formula indicates standards of consistency in willing and Guyer rightly refers to this notion of consistency in his view of humanity though his identification of “humanity” with freedom has some difficulties.

Let’s regress a little, as Guyer attempted to do. Guyer indicated the possession of freedom is, in an important way, a condition of the possibility of applicability of universal law. However, if the notion of “humanity” means at least possession of rational willing, there must be something more involved in this than an exercise of freedom since there appears nothing problematic in itself in freely choosing to do that which is not expressive of respect for others or for the “humanity” in them. So if “humanity” indicates a kind of capacity that Rawls and others can relate to moral personality and it references a kind of purposive structure, and it is connected to the formulation of universal laws, then it follows that the capacity it describes must be more than simple freedom of choice or purposiveness in general. Rather than adopt these minimal characterisations it is more apt to describe the characteristic of “humanity” as that which makes it possible for us to be moral (Ak. 4: 435). This would be the ability not simply to act in accordance with

maxims that conform to the requirement of universal law but to actually be capable of setting universal laws for oneself. In other words, the capacity that Kant terms “humanity” would be equivalent to the ability to will autonomously. This is why Kant can give the formula of humanity as effectively equivalent to the formula of universal autonomy as he does in the following statement:

“. . . the principle: so act with reference to every rational being (to yourself and others) that in your maxim it counts at the same time as an end in itself, is fundamentally one and the same as the principle: act on a maxim that at the same time contains in itself its own universal validity for every rational being. For to say that in the use of means to every end I ought to limit my maxim to the condition of its universal validity, as a law for every subject, is tantamount to saying that the subject of ends, i.e. the rational being itself, must be made the foundation of all maxims of actions, never merely as a means, but as the supreme limiting condition in the use of all means, i.e. always at the same time as an end.” (Ak. 4: 438)

Now, some of the sense of this statement still needs teasing out, particularly in relation to the claim that is made concerning the rational being as the “supreme limiting condition” in the use of all means. This point was one I noted earlier was made by Korsgaard and connected by her also to the sense that the capacity of “humanity” was capable of greater or lesser realisation. If we adopt the view that “humanity” and “autonomy” are effectively equivalent then it follows that it is the capacity to be able to set universal moral ends that is the supreme limiting condition of all other ends but it also follows from this account that there are varied degrees of “humanity” expressed both between people and, in the same people, at different points. Now that we have a description at least of what the capacity for “humanity” consists in we must use this in order to help

us to determine the general interpretation of the Formula of Humanity, since, as we shall see, there is considerable controversy concerning what this formula in a general sense involves, not just what the capacity referred to by it as “humanity” consists in.

Formal and Substantive Views of the Formula of Humanity

Essentially there are two types of general reading of the Formula of Humanity and they can be divided by stating that according to one, the Formula can be construed in a generally formal way, whilst according to the other it introduces a substantive value. The first, formal, reading of the formula is the one that was, until recently, the standard or received view of it whilst the second, substantive, reading is a more recent one that has arisen partly as a consequence of thinking of Kant’s formulas of universal law as inadequate on their own and as requiring supplementation of a sort that involves recourse to a view of value that the Formula of Humanity is taken to have enabled. However, whilst this general contrast is one that is capable of being represented relatively clearly, the details of the different types of reading under each heading and the motivations for adopting one reading or the other, are correspondingly much more complex.

If a formal view of the Formula of Humanity is adopted it is rarely due simply to attention to the formula itself. It is, rather, a consequence of a general methodological view of the shape of ethical reasoning. The basic form of this view is found in a remark Kant makes in the *Critique of Practical Reason* which is applied not just to the method given in this work but also used as a model for how to interpret the *Groundwork*. This remark concerns the relationship Kant states there is between concern for the moral law, on the one hand, and accounts of the good, on the other. Kant writes here of what he terms “the paradox of method” by which the *Critique of Practical Reason* is governed and this paradox is

that “*the concept of good and evil must not be determined before the moral law*” but rather only “*after it and by means of it*” (Ak. 5: 63). This “paradox” of method involves placing the notion of the moral law over and above any account we might have of the good and ensures that no specific notion of the good is initially assumed in ethical theory but rather that any account of what is taken to be good must rather be based on the manner by which the process of universalization itself proceeds. The idea that is derived from this claim by those who uphold formal views of Kant’s procedure (and hence of the Formula of Humanity) is known as a belief in the priority of the right over the good. Samuel Freeman states that this notion of the priority of the right “concerns the structure of the practical reasoning of moral agents” and describes “the content internal to principles of right”.⁹

The basic idea is that principles of right are built into the structure of formal law-giving itself and that conceptions of the good are defined and limited by reference to this. We have seen some grounds for thinking that the reasoning behind Kant’s use of the Formula of Humanity can be described in a manner that fits this conception. However, the basic objection that can be put to an approach of this sort, consists in focusing on the way in which the Formula of Humanity introduces reference to the notion of an *end* since it is commonly assumed that upholders of the priority of the right are engaged in an abstraction from considerations of “ends” due to their methodological subordination of considerations of the good to the right.¹⁰

⁹ Samuel Freeman (1994) “Utilitarianism, Deontology, and the Priority of Right” in S. Freeman (2007) *Justice and the Social Contract: Essays on Rawlsian Political Philosophy* (Oxford University Press: Oxford and New York), pp. 66-7.

¹⁰ In a sense this is due to two points, both an assimilation of the priority of the right to “deontology” and a view of the latter as requiring absence of any reference to value. Freeman’s article disputes the assimilation and presents “deontology” as requiring not abstraction from value but, instead, a pluralist view of the good. For the view that “deontology” involves abstraction from value see Barbara Herman (1993) *The Practice of Moral Judgment* (Harvard University Press: Cambridge, Mass and London), Chapter 10.

I am going to consider three forms of “substantive” readings of the Formula of Humanity before proceeding to test whether the general methodological considerations they advance truly mandate departure from the idea of the priority of the right and, if they do, whether such departure from the priority of the right is consistent with Kantian principles. Part of the problem in assessing such views, however, concerns what they mean when they claim, as they generally do, that Kant affirms a “substantive value” in the Formula of Humanity. What is meant by this notion of “substantive value”? Here Christine Korsgaard is of some help since she distinguishes between two types of contrast that are often at work in considerations of stating something either to be “good” or to have “value”. On the one hand, we may be claiming that what is so “good” or of “value” has such a characteristic “intrinsically” or, conversely, that it derives its value from something external to itself, and thus has an “extrinsic” form of goodness. This contrast is to be thought as distinct from one where we claim that something has “value” in the sense that we don’t regard it as good *for* something but rather as being a “final” good that is prized for its own sake. Notice that taking something to be “intrinsically” good is not necessarily equivalent to saying that it is a good of a “final” sort. For example, we might hold the view that inasmuch as art-works have value that this value is intrinsic to them and yet still not regard art-works as having a value that can be prized above all else as we would think if we took them to be “final” goods. Goods that are “final” in this sense are thus not comparable with anything further and in being so incomparable have a status that raises them above all else. “Final” goods thus construed are to be contrasted, however, not with “extrinsic” goods but rather with “instrumental” goods, which latter are *good for* something else.¹¹

¹¹ These distinctions are made in some detail and connected to an important contrast between Kant and intuitionist views in C. Korsgaard (1983) “Two Distinctions in Goodness” in C. Korsgaard (1996) pp. 249-74.

When this contrast between two distinctions is made it should be clear that the way Kant speaks of the characteristic of “humanity” in his account of the formula is primarily in terms of thinking of it as a “final” good which is explicitly contrasted with instrumental goods. The contrast between “intrinsic” and “extrinsic”, by contrast, is not predominant in Kant’s account of the Formula of Humanity, presumably because, whilst something having an intrinsic kind of goodness tells one of the *source* of its value, it does not, just by means of this, indicate the kind of *priority* to attach to the value in question. However, it is possible to think of a final end that is extrinsic where it is final in the sense that we do not require some further reason for desiring it and yet find its good to be an extrinsic one. Such might be “happiness” for example, whose goodness is final since if something makes us happy that seems sufficient reason on itself for desiring it and yet happiness is itself a state that has an extrinsic relationship to goods such that they are objects of acquisition in a “material” sense. Hence, it would seem that “humanity” is not merely a final end but also an intrinsically final end and that it is its possession of both these characteristics of goodness that makes it specifically valuable. So, whether the reading offered of the Formula of Humanity is formal or “substantive”, it should reflect this dual quality of the goodness that is attributed to “humanity” by Kant.

Not only should a “substantive” reading of the Formula of Humanity indicate how the value it purports to find in the formula is a “final” end in the sense of a basic and overriding reason that has decisive weight in deliberation, it should also help us to sort out the way in which the two requirements I mentioned in opening as emerging from the formula are to be weighted. These requirements are, to remind you, that we treat persons with respect and that we always treat them, when we are relating to them as means, as being at the same time ends in themselves. I want to suggest that there are two tiers of consideration in Kant’s account and that it is important to distinguish them but I will not justify this yet, merely mention it as a constraint on

distinct readings of the formula and, most especially, for reasons that will become swiftly clear, on “substantive” readings of it.

The first type of “substantive” reading I want to mention comes from Allen Wood. As we’ve already seen Wood views the capacity of humanity to reside in rational nature but appears to include in this value two distinct conceptions, both the minimal view of this nature as resting only in purposiveness in general and the more extended view of setting not just ends *simpliciter* but “good” ends. However, if the value that is revealed in the Formula of Humanity has a structure that is significantly distinct from that prized in formal readings, this value cannot just reside in the capacity to formulate and follow universal law, as otherwise Wood’s account turns out not to be “substantive” after all. Wood’s “final” value does however reside in rational nature since he says that the “goodness of any other end must somehow be grounded in this value”.¹² Since this is so it can only really be a question of how the “requirements” that arise from the value of humanity so considered that gives Wood’s view a way of departing from the classic formal approach. The means by which Wood marks this departure is through taking the respect requirement to be the one that must be given most weight. This is stated in the clearest way when he argues that the Formula of Humanity is a vehicle for Kantian social criticism as when he writes: “when a social order treats some people better and some worse in ways that they themselves regard as essential to their self-worth, there is a presumption, based on FH itself, that this social order fails to respect the humanity of those who receive worse treatment”.¹³

Regardless of the possible results of this way of construing the Formula of Humanity I want to suggest that there are two defects with Wood’s account of it. Firstly,

¹² Wood, *op.cit.*, p. 92.

¹³ Allen Wood (1995) “Humanity As End in Itself” in Paul Guyer (ed.) (1998) *Kant’s Groundwork of the Metaphysics of Morals: Critical Essays* (Rowman & Littlefield: Lanham and Oxford), p. 183.

whilst taking the capacity of “humanity” correctly to be a “final” and over-riding requirement in ethical deliberation Wood gives little attention to how such deliberation is to take place other than to state that what the formula demands of us is that we, as he puts it, “*express* due respect for the worth of humanity” in our actions.¹⁴ The view that the shape of deliberation should be in terms of how actions express “respect” both requires us to take the respect requirement as prior to the requirement for taking humanity as an end in itself and indicates a preference for a conception of “expression” that highlights a consideration for attitudes that is not evidently of a piece with the priority of the right. Finally, in his recipe for social criticism as arising from his view of the formula Wood opens the door to relative considerations of value given that his version of the respect requirement turns on a pre-given view of what agents themselves hold to be their due prior to consideration of its relationship to the law.¹⁵

By contrast to Wood, Guyer, as we have already seen, takes the capacity of humanity to reside principally in freedom though, in highlighting freedom rather than autonomy, he leaves himself open to an obvious retort. However, Guyer’s conception that the value of “humanity” resides in freedom, does involve an appeal to a procedure of universalization and in this respect is preferable in its congruence with Kantian procedure, to Wood’s account of the respect requirement. But Guyer is clear that the appeal to freedom is one that is meant to provide us with a value that leads us to adopt the procedure of universalization, rather than being something whose internal consistency is tested by means of the procedure of universalization. Not only is this so, but, in making this argument he articulates a conception of freedom that is illuminating in terms of the consequences

¹⁴ Wood (1995) p. 177.

¹⁵ In conceding to subjectivism in this way Wood essentially equates the desire for something with its intrinsic goodness rather than understanding that desire is only extrinsically related to goodness as Kant argues.

that arise from his methodological claim that adherence to the law emerges from commitment to a pre-existent value:

On this approach there is no initial assumption of the moral value of adherence to the moral law for its own sake; rather, freedom per se is intrinsically valuable, but the introduction of consistency both within one's own choices of ends and the ends of oneself and others by means of compliance with the requirement of universalizability is argued to be necessary in order to maximize the exercise of this freedom and thus maximally realize its potential intrinsic value. On this account, freedom is the absolute value and adherence to the law is the condition necessary for the maximal realization of this value rather than a part of its very concept.¹⁶

This type of “substantive” conception does involve a very marked departure in methodology from the traditional formal view. If the antecedent value of freedom is what is reflected in the Formula of Humanity and the conception of its self-consistency is further determined by a process of maximisation of this value then essentially Guyer's account of the procedure of Kantian ethics verges on consequentialism, a result that is distinct from Korsgaard's description of levels of realization of humanity but which clearly emerges from a view of a similar type. Further, Guyer takes the value that is antecedent to the law to be an “intrinsic” rather than “final” value and yet views the law, by contrast to the “intrinsic” end, as having an instrumental rather than “extrinsic” point. If we take humanity to have this “intrinsic” value there is a basis for taking other “intrinsic” values to be higher than freedom itself since we have not determined freedom's value to be final and, further, if the means by which freedom is maximally realized are through adherence to a procedure of universal self-consistency the process of

¹⁶ Paul Guyer (1991) “Kant's Morality of Law and Morality of Freedom” in P. Guyer (2000) *Kant on Freedom, Law, and Happiness* (Cambridge University Press: Cambridge and New York), p. 156.

evaluation still requires appeal to a formal notion of procedure. This point about reference to formal processes is explicitly stated by Guyer who adds that the formula of universal law functions as the “decision procedure by which any agent would test his proposed maxims of action”,¹⁷ which ensures that despite his having found a *value* in the Formula of Humanity which the formal process has to realize by maximising, that the procedure of decision concerning the application of this, still adheres to a formal structure.

If Wood’s departure from the traditional reading thus appeared to displace reference to universal law and court a form of relativism in its construal of the respect requirement, Guyer, by contrast, incorporates both reference to a value and a procedure of maximisation into an account that, despite its novelty, resolves essentially into an appeal for a formal procedure of consistent universalization that undercuts its apparent reversal of the priority of the right. Neither of these forms of “substantive” reading appear, then, to succeed and, in Guyer’s case, there seems to be a methodological muddle since decision procedures of universal consistency are indicated to harmonize in some odd manner with reference to maximisation notions.

By contrast to both these approaches, and, much more boldly, David Cummiskey advances a frankly consequentialist view both of Kantian ethics in general and of the Formula of Humanity in particular. Cummiskey does not think that Kant’s view is consequentialist in the sense that the justificatory arguments Kant gives for the moral law can or should be construed in a consequentialist way but he does think that the normative theory that emerges from consideration of the Formula of Humanity has a consequentialist structure. In adopting this view Cummiskey is following the same direction of thought as Guyer but taking it one step further. Cummiskey views the Formula of Humanity as providing us with an “internal perspective” on agency such

¹⁷ Guyer, *ibid*, p. 161.

that the “objective end of rational action must be a necessary presupposition of rational agency itself” so that this objective end is a final end for Kant.¹⁸ This picture allows him to view the general justification of Kant’s reference to humanity to be self-reflexively available to the agent when morally deliberating and to take this element of Kant’s picture to have a formal structure.

Despite Cummiskey’s recognition of this formal structure of self-reflexive deliberation he nonetheless takes Kant’s normative theory to be one that provides, as he puts it, “an objective end or moral goal for a consequentialist principle of right” (*KC* 85) and he justifies this claim by means of a five-step argument he terms the “equivalence argument”. The steps of this argument have the following form:

1. If rational nature is an end-in-itself that has absolute value then it is a goal we are required to promote.
2. Rational nature is an agent-neutral value and hence has equal standing in distinct agents.¹⁹
3. On the basis of 1) and 2) I should seek to promote the necessary conditions of rational agency, including reflective choice and rationally chosen ends, regardless of agent-relative considerations.
4. There are no additional constraints on the maximization of the good than recognition of the constraints of promotion of the ends of rational nature.
5. I am rationally required to *maximally promote* the objectively valid ends of morality.²⁰

¹⁸ David Cummiskey (1996) *Kantian Consequentialism* (Oxford University Press: Oxford and New York), p. 65, hereafter referred to in the text and notes as *KC*.

¹⁹ This claim is derived by Cummiskey from Thomas Nagel (1986) *The View From Nowhere* (Oxford University Press: Oxford and New York), Chapter 6. It is, also, however, articulated by Christine Korsgaard (1993) “The reasons we can share: An attack on the distinction between agent-relative and agent-neutral values” in C. Korsgaard (1996), which, as its title suggests, seeks to undermine the distinction it requires. Cummiskey’s acknowledgement of this point ensures, however, that his view is not subject to the objection made above against Allen Wood.

²⁰ This argument, the key to the whole of Cummiskey’s book, is presented in compressed form on pp. 87-91 of *KC*.

With this argument Cummiskey advances a distinctive “substantive” view of the Formula of Humanity that justifies a consequentialist theory in the sense that there is a good that is taken to be a central value and the structure of the theory tells us to maximise this good. The obvious retort to Cummiskey is that this theory does not respect the Kantian requirement that the notion of “humanity” act as a restraint upon permissible actions but Cummiskey rejects this response stating that on this view “rational nature as such and the legitimate ends of others are indeed constraints” on subjective ends (*KC* 89). The picture produced has a great deal of structural similarity to Guyer’s view but, unlike Guyer, in taking rational nature as a goal that is capable of maximizing, it results in building the conception of maximisation directly into the means by which formal consistency of the end aimed at is assessed (rather than being, as with Guyer, a process whereby formality of consistency is intended to be merely self-regulated).

The fifth and final part of the “equivalence argument” produces the sought consequentialist end but the move from the second to the third premise, which is crucial for this result, involves a move from a general to a specific kind of equivalence. In the “general” claim of the second premise the move is made that in choosing rational ends I choose not merely universally in accord with notions of rational consistency but that I should do so in such a way that abstraction is made away from any specific ends I possess. Whilst this involves recognition of the importance of the reference to necessary ends in Kant’s thought it appears to evacuate any space for taking any ends that are specific to me seriously regardless of their compatibility with a more formal construal of the humanity test. After making this point Cummiskey moves to the more specific follow-on that rationally chosen ends of others are to be taken as indistinguishable from my own ends. This has the merit of interpreting the mere means requirement to be primary over the respect requirement and thus ruling out any moves of the

type Wood wishes to make, although, as I shall say soon, it does so by rendering the mere means requirement one that lacks Kantian content. However, there is a key problem with Cummiskey's argument from premise 2 to premise 3, which, is, that unless an additional requirement is added between them we will have a *non sequitur*. This point is well made by Daniel Weinstock who writes:

It simply does not follow from the fact that I should view the rational nature of all other agents as being of equal value to my own that I should view their rationally chosen ends as ones that I ought to pursue to the same extent and in the same manner as I pursue my own ends. At the best, this conclusion only follows from the general formulation of the equivalence argument if an implicit premise is assumed, namely, that it does not make a difference to the value realized through the attainment of an agent's rationally chosen ends, whether these ends are attained through the agent's own activity or through the agency of another person. Let me call this suppressed premise the *indifference claim*.²¹

Only if the "indifference claim" is inserted between premises 2 and 3 does the transition from the general equivalence claim to the specific one really function and the problem with the "indifference claim" is that it clearly compromises respect for autonomy and does more than over-ride the respect requirement, it prevents recognition of it as having any force at all. In thus compromising reference to autonomy Cummiskey's account of the requirements emergent from his form of "substantive" interpretation of the "humanity" formula ceases to have serious Kantian credentials.

²¹ Daniel Weinstock (2000) "Critical Notice of David Cummiskey *Kantian Consequentialism*", *Canadian Journal of Philosophy*, Vol. 30, No. 2, pp. 322-3.

Value and the Structure of Kant's Moral Theory

Since all the formulations of “substantive” conceptions of the Formula of Humanity considered appear to run into serious problems either with recognition of the requirements that Kant assumes will emerge from the formula or with comprehending serious congruence between the formula and Kant’s accounts of universalization, it is worth, in conclusion, detailing some of the ways value is structured by Kant’s moral theory in order to show the ways in which a reading of the formula should be construed if it is to be consistent with the structural limits of Kant’s theory. The result of this will be to suggest that whilst a formal account of the formula is to be preferred that there are ways to show consistency of such an account with the stress on ends that “substantive” readings have been aiming for.

In opposition to the general view of according value to freedom that Guyer states, I advocated, earlier, that we view instead the Formula of Humanity as essentially equivalent in status to the Formula of Autonomy, which asks us to think of universal laws not merely as something that we are capable of construing our maxims as harmonious with, but also as something we choose to will as a law for ourselves. The active choice of the law as the guide for deliberation is, I am suggesting, the exercise of the capacity called “humanity”. If that is right then the value that “substantive” readings of the Formula of Humanity are seeking should not be taken as something independent of or antecedent to the law. The reason why it appears that a choice has to be made for a value that is construed as importantly distinct from the law is due to the way the Formula of Humanity builds in a clear reference to “ends” that is not specified in the formulas of universal law. When Kant applies the formulas of universal law, however, it is to maxims that already contain stated in them reference to an end of some sort or other. This is the means by which the categorical imperative stands in an important relation to the hypothetical imperative but, whilst

the hypothetical imperative gives the rule for instrumental setting of ends, the categorical imperative, by contrast, is intended to state a relation to ends, that are viewed not instrumentally but finally. So, the end specified in the Formula of Humanity is meant to be distinct from any other type of end, and the problem with “substantive” readings of the Formula of Humanity is that this distinction between the end that is an end-in-itself, and ends that are only instrumental, is not clearly enough picked out by them.

The reason why “substantive” interpretations of the Formula of Humanity have proved incapable of respecting the specific character of the end stated in the formula is due to departure, I want to argue, from the priority of the right. If we view the right as prior to the good so that any notion of the good concerns its manner of conformity to law then the procedure of consistent universalization will determine the worth of any ends that are aimed at in moral terms. Now, if the capacity of “humanity” is, as I have argued, the capacity to freely set ourselves universal self-consistent ends, then, it is this capacity that governs the whole procedure of formal universalization, and, hence, respecting the end set in the formula cannot involve viewing the Formula of Humanity as giving us warrant for a different procedure from that which we adopt in the formulas of universal law.

So if there are values that are loaded into the process of self-consistent universalization then these values can be no other than the ends of pure practical reason itself. So the basic value that would govern the procedure of universalization of the Formula of Humanity should be reflective of the priority of the right over the good. If this is the basic value, this does not prevent it from being the case that there is also what we might term a “higher-order” value, that determines the procedure of priorities that any given subject would specify as central, and this higher-order value is essentially what we term the respect requirement. On the one hand, then, the Formula of Humanity states as cardinal and basic to its operation the mere means requirement since the capacity of “humanity” is an end in itself and this

capacity in general terms cannot be regarded as something that is fungible. However there is also a requirement that relates to the presence of the capacity for “humanity” in ourselves and others that indicates that all persons are worthy of respect precisely as capable of exercise of their capacity for “humanity” and whose dignity and worth resides in their ability to exercise this capacity above all else.

Departure from the priority of the right typically produces a teleological value theory, which means, as Rawls defines it, that “the good is defined independently from the right, and then the right is defined as that which maximises the good”.²² The classic form of such a theory is consequentialism and Cummiskey’s equivalence argument indicates the move that such a view requires which is a form of abstraction from ends not simply due to their contingent character as Kant might be thought to require in consideration of the moral worth of maxims, but also in abstraction from concern for the humanity present in any given individual. In requiring acceptance of this condition, that Weinstock aptly termed an “indifference claim”, the consequentialist theory indicates structural lack of concern with respect for persons.²³

The reason why it matters that consequentialist theory has this structure is because it is led to it by a view of morality that cannot meet the requirements that Kant finds necessary. Essentially, in taking value to be primary in moral theory, the consequentialist position is led to adopt maximisation as an understanding of the means to realise the value isolated, as the procedure of maximisation is simply assumed to have no elements in it that could violate the value in question. This is well revealed in Cummiskey’s suggestion that it is only a “confusion” to think that his consequentialist reading of the Formula of Humanity is not one that structurally

²² John Rawls (1971) *A Theory of Justice* (Oxford University Press: Oxford and New York), pp. 21-2.

²³ Rawls famously specified this point in relation to utilitarianism as requiring that utilitarians not take seriously the separation of persons. The point here is that this characteristic is generic for consequentialists, not simply an indication of a feature of its most prominent version.

respects the constraints imposed by the requirement of respect for humanity. The procedure is one whereby, persons are required to relate to the quality of “humanity”, as something that is essentially ontologically distinct from its possession by persons, so that persons can be by-passed in terms of its realisation. On this conception there is no specific quality of worth in the moral growth of any person that could not be matched equally well by a general impersonal growth of the capacity of humanity in general. Due to this procedure it follows that persons are effectively fungible, in direct contravention of the mere means requirement, and as a consequence of rejection of the respect requirement.

The basic thing that goes wrong in such a procedure is the failure to view certain ends as requiring a specific relationship that is distinct from all other ends. With hypothetical ends, it can be right to think that the way to realise them is not by directly aiming at them. For example, economists speak of the “paradox of thrift” that arises when everyone saves, producing a situation where aggregate demand is reduced and thus the net worth of savings falls. This illustrates that hypothetical imperatives may require, for their realisation, action that appears to aim at the opposite to what one requires. By contrast, the end that is stated in the Formula of Humanity is a categorical end, which is intrinsically compromised if one acts against it, as one would if the neglect of the respect requirement was an essential element of one’s procedure. It is precisely due to not viewing “humanity” as a good that can be related to after the manner of instrumental reasoning but as requiring instead a procedure that reflects its status as a good that is both intrinsic and final that recognition of this end requires adoption of the priority of the right. Abandonment of the priority of the right is equivalent not merely to abandonment of Kantian ethics but also of a concern with ends that are ethically intrinsically unable to be treated as mere means to something further.