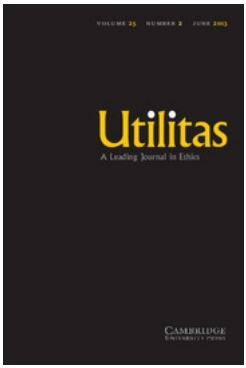
Utilitas
http://journals.cambridge.org/UTI

Additional services for ***Utilitas:***

Email alerts: Click here
Subscriptions: Click here
Commercial reprints: Click here
Terms of use : Click here



The Common Structure of Kantianism and Act-Utilitarianism

CHRISTOPHER WOODARD

Utilitas / Volume 25 / Issue 02 / June 2013, pp 246 - 265
DOI: 10.1017/S0953820812000489, Published online: 09 May 2013

Link to this article: http://journals.cambridge.org/abstract_S0953820812000489

How to cite this article:
CHRISTOPHER WOODARD (2013). The Common Structure of Kantianism and Act-Utilitarianism. Utilitas, 25, pp 246-265 doi:10.1017/S0953820812000489

Request Permissions : Click here

CAMBRIDGE JOURNALS

Downloaded from http://journals.cambridge.org/UTI, IP address: 78.86.85.214 on 03 Jul 2013

The Common Structure of Kantianism and Act-Utilitarianism

CHRISTOPHER WOODARD
University of Nottingham

This article proposes a way of understanding Kantianism, act-utilitarianism and some other important ethical theories according to which they are all versions of the same kind of theory, sharing a common structure. I argue that this is a profitable way to understand the theories discussed. It is charitable to the theories concerned; it emphasizes the common ground between them; it gives us insights into the differences between them; and it provides a method for generating new ethical theories worth studying. The article briefly discusses the relationship between these ideas and some other recent proposals that emphasize the common ground between Kantianism and versions of consequentialism.

There is a way of understanding Kantianism and act-utilitarianism according to which they are both theories of pattern-based reasons, and thereby share a common structure. This article explains this way of understanding them, and what it can contribute to our understanding of these and other important views.

I cannot hope to do justice to the richness of Kant's ethics, and I will not be engaged in exegesis of his texts. Instead I will discuss 'Kantianism', an ethical view to be defined in section I. It is, I hope, sufficiently close to Kant's ethics for my claims about it to be of interest. In section II I will explain how this view can be understood as a theory of pattern-based reasons. In section III I will discuss act-utilitarianism, explaining how it too can be understood as a theory of pattern-based reasons, before explaining in section IV what can be gained by understanding these views in this way.

Others have argued that Kantian premises support consequentialist conclusions.[1] Understandably, these proposals are very controversial.[2] But my proposal is not of this sort. Instead, I will be arguing that

[1] H. Sidgwick, *The Methods of Ethics*, 7th edn. (London, 1907), p. xxi; D. Cummiskey, *Kantian Consequentialism* (Oxford, 1996), p. 4; R. M. Hare, *Sorting Out Ethics* (Oxford, 2000), pp. 153–4; S. Kagan, 'Kantianism for Consequentialists', *Groundwork for the Metaphysics of Morals*, ed. A. Wood (New Haven, CT, 2002), pp. 111–56, at 150; D. Parfit, *On What Matters*, ed. S. Scheffler, 2 vols. (Oxford, 2011), vol. 1, pt. 3.

[2] R. Dean, *The Value of Humanity in Kant's Moral Theory* (Oxford, 2006), pp. 165–74; S. Morgan, 'Can There Be a Kantian Consequentialism?', *Ratio* 22 (2009), pp. 19–40; M. Otsuka, 'The Kantian Argument for Consequentialism', *Ratio* 22 (2009), pp. 41–58; J. Ross, 'Should Kantians be Consequentialists?', *Ratio* 22 (2009), pp. 126–35.

© Cambridge University Press 2013
Utilitas Vol. 25, No. 2, June 2013
doi:10.1017/S0953820812000489

there is a symmetrical relationship between Kantianism and act-utilitarianism: they are both versions of the same sort of theory, but the premises of neither support the conclusions of the other. In particular, I will not be trying to recruit Kantians to the consequentialist cause.

Some others have made similar claims about shared structure. According to the 'consequentializing idea', every normative theory can be represented as a version of act-consequentialism.[3] This claim entails that Kantianism and act-utilitarianism share the structure given by act-consequentialism; differences between them are the result of differences in their theories of value. A different proposal has been made by Michael Ridge. He has shown that Kantian verdicts can be generated by a theory that employs an *agent-neutral* theory of value, provided it incorporates a non-standard theory of rightness and certain other assumptions.[4] Both of these proposals are ways of understanding Kantianism as sharing a structure with act-utilitarianism, and in section V I will briefly discuss the relationship between them and the proposal made here.

I. KANTIANISM

Every rational agent is capable of acting, on every occasion, as a member of the Kingdom of Ends. Acting in this way involves acting on maxims that respect the humanity of other rational agents by treating them as ends in themselves. Some maxims do not respect the humanity of other rational agents. These maxims cannot be willed as universal laws. This might be because there is a conceptual incoherence in the idea that they serve as universal laws governing the behaviour of all rational agents. Alternatively, it might be because a world governed by them as universal laws would be significantly worse than some other practically possible world. Other maxims can be willed as universal laws, because they have neither of these flaws. Whenever a rational agent acts on one of these maxims, she thereby acts as a member of the Kingdom of Ends and also succeeds in respecting the humanity of other rational agents.

We can imagine that every rational agent always acts on one of the permitted maxims. If that were so, the Kingdom of Ends would be realized. Every rational agent would always respect the humanity of

[3] D. W. Portmore, 'Consequentializing Moral Theories', *Pacific Philosophical Quarterly* 88 (2007), pp. 39–73.

[4] M. Ridge, 'Consequentialist Kantianism', *Philosophical Perspectives* 23 (2009), pp. 421–38.

every other rational agent, by acting only on maxims that could be willed as universal laws. In so doing, every agent would be acting at all times as a member of the Kingdom of Ends.

This possibility could be realized in different ways. The permissible maxims do not specify exactly one action for each circumstance. Instead, they rule out some kinds of action altogether, and leave room for choice between others. However, there is a further constraint on the choices made by rational agents acting as members of the Kingdom of Ends. Some *sequences* of actions, composed only of actions that are permitted in isolation, are not themselves permitted. For example, it is permitted, on any given occasion, for a rational agent not to take an opportunity to develop her talents. However, it is not permitted for any rational agent always to fail to take such opportunities. She could not rationally will living her whole life without developing her talents, since this would be to fail to respect her own humanity. We can summarize these points by saying that, if the Kingdom of Ends were realized, certain actions would never be performed at all. These actions are prohibited by perfect duties. Other actions could be performed if the Kingdom of Ends were realized, though certain sequences of them could not be realized. These actions are governed by imperfect duties, which prohibit some sequences but not each isolated action.

Each of the ways in which the Kingdom of Ends could be realized is a pattern of action. Each such pattern consists of all the actions performed by all rational agents in that realization. Each pattern thus consists only of permitted actions and permitted sequences of actions. In each such pattern, every rational agent respects the humanity of every other rational agent by acting only on maxims that could be willed as universal laws. Each such pattern of action would have great value, since it would be the practical expression of unfailing good will, and would consist only of actions that respect the humanity of rational agents. We can call any such pattern of action a 'Kingdom of Ends pattern', or *KE pattern* for short.

What ought any agent to do on any occasion? She ought to act in some way that realizes, on that occasion, a part of a KE pattern. In doing this, she will be acting in accordance with the perfect duties. Additionally, she ought to live her whole life in a way that realizes part of a KE pattern. In doing this, she will be acting in accordance with the imperfect duties. Importantly, she is not required to maximize the extent to which any KE pattern is, on the whole, realized. She is required only to perform her parts of some KE pattern.

What reasons does any agent have on any occasion? Possibly, she might have reasons that are not associated with the prospect of

realizing part of some KE pattern.[5] But if she does have such reasons, and they ever conflict with the reasons she has to realize part of some KE pattern, it is the reasons to realize part of a KE pattern that are decisive. They determine what she ought to do, though they may leave scope for the operation of other reasons within the constraints on action they provide.

The ethical view that I have described is what I will call 'Kantianism'. This view, I will claim, is of the same basic kind as act-utilitarianism. I do not claim that it is Kant's own view exactly. However, the interest of my claims will depend partly on the degree to which it resembles Kant's own view.

II. PATTERN-BASED REASONS

Kantianism claims of every rational agent in every circumstance that she ought to perform part of a KE pattern, because she has a decisive reason to do so. This claim employs the idea of *pattern-based reasons*. These are normative reasons of the following form: the fact that an agent X could perform some action A in circumstances C, and that A would be part of some favoured pattern of action P, is a reason for X to perform A in C. In the case of Kantianism, every KE pattern is favoured, and it is claimed not only that in all circumstances the agent has a reason to realize part of a KE pattern, but also that such reasons are decisive in the sense that no other reasons she may have could defeat them. As a result, every agent in every circumstance ought to realize part of a KE pattern.

Contrast pattern-based reasons with instrumental reasons. An instrumental reason has the following form: the fact that an agent X could perform some action A in circumstances C, and that A would cause some favoured outcome O, is a reason for X to perform A in C. An instrumental reason is a reason that depends on a *causal* relationship between the action for which it is a reason and something that is favoured (an outcome). A pattern-based reason is a reason that depends on a *parthood* relationship between the action for which it is a reason and something that is favoured (a pattern of action).[6]

One good question is whether the idea of pattern-based reasons makes sense. Most of us agree that there are reasons to cause something good that depend on the causal relationship between the action and

[5] This depends on how finely KE patterns are individuated. If anything that could provide a reason is also sufficient to distinguish one KE pattern from another, there will be no reasons for action that are not associated with KE patterns, according to Kantianism.

[6] Compare Nozick on the idea of symbolic value. R. Nozick, *The Nature of Rationality* (Princeton, 1993), pp. 26–35, 41–50.

the good. It is perhaps less clear whether the parthood relation also supports reasons. This is not centrally at issue in this article, since I am not here defending Kantianism or the idea of pattern-based reasons more generally.[7] However, it is important to my claims here that Kantianism as I have defined it retains at least some of the appeal of Kant's ethics. So I will say a few things in favour of the idea that the parthood relation supports reasons. The three points that follow are designed to make the idea seem minimally coherent, rather than to serve as a full defence of it.

One point is that the idea of pattern-based reasons offers the most straightforward interpretation of some of the things we say when trying to justify actions. Why did the keen gardener comply with the hosepipe ban when he could certainly have got away with flouting it? 'I was playing my part in tackling the water shortage.' Why does the vegetarian refuse to eat animals that are already dead? 'I don't want to participate in the practice of killing animals for food.' Why does the elector vote? 'I want to elect the right government.' In each of these cases, the most straightforward interpretation of the putative justification is that a pattern-based reason is being adverted to: refraining from using the hosepipe, refraining from eating meat and voting in the election are each parts of evidently favoured patterns of action, and none of them is sufficient to cause the favoured pattern to be realized in full or to bring about the outcome that makes the pattern favoured. In each case, it is the parthood relation (between the action and the favoured pattern) rather than a causal relation that appears to be the putative source of reasons. This is evidence that in ordinary moral thought we take the idea of pattern-based reasons seriously.

A second point is that some other views in moral philosophy appear to use the idea of pattern-based reasons. I will give further examples later, but for now consider rule-consequentialism. Versions of this theory that evaluate sets of rules according to the consequences of universal (or near-universal) compliance with them transparently make use of the idea of pattern-based reasons, since the consequences of compliance *are* the consequences of a favoured pattern of action.[8] These theories claim that each of us ought to perform actions only if they are permitted by the best set of rules. They thus claim that we have decisive reasons to realize parts of favoured patterns of action (just because they are parts of favoured patterns of action). Just as ordinary moral thought

[7] For a defence of the idea, see C. Woodard, *Reasons, Patterns, and Cooperation* (New York, 2008).

[8] C. Woodard, 'A New Argument against Rule Consequentialism', *Ethical Theory and Moral Practice* 11 (2008), pp. 247–61.

appears to make use of the idea of pattern-based reasons, so does moral philosophy.

Finally, we should be careful to distinguish between objections that can be made to any particular theory of pattern-based reasons, and objections to the coherence of the idea of pattern-based reasons itself. For example, philosophers often worry about the idea that there could be a reason to realize part of a favoured pattern in cases where it is certain that the remainder of the pattern will not be realized. In some of these cases, realizing one part alone would have bad consequences. To take an example that is especially germane to the present discussion, consider the suggestion of Kantians that we should not lie. This is a perfect duty, since no maxim permitting lying could be willed as a universal law, on pain of incoherence (according to Kantianism). Hence no KE pattern contains any instance of lying, and no agent in any circumstance is permitted to lie. The present objection is this: where we know for sure that others will not realize their parts of a KE pattern, and that realizing one's own part will have terrible consequences (as in the case of the murderer at the door), it is incredible to claim that there is a reason – worse, a decisive reason – to play one's part. For example, Christine Korsgaard says: 'in cases such as that of the murderer at the door it seems grotesque simply to say that I have done my part by telling the truth and the bad results are not my responsibility'.[9] Whether or not the exceptionless prohibition of lying can be defended, the present point is that this is an objection to one particular use of the idea of pattern-based reasons, and should not be treated as an objection to the idea itself. The idea of pattern-based reasons does not itself imply the existence of any exceptionless prohibitions. Moreover, interpreting Kantianism as a theory of pattern-based reasons does nothing to make the problem of the murderer at the door worse.

Thus I will assume in the remainder of this article that the idea of pattern-based reasons makes sense, and that interpreting a moral theory as employing it is not obviously uncharitable. Now let us note some ways in which theories of pattern-based reasons can differ from each other. I will identify three dimensions of variation in total.

Obviously, one dimension in which they can differ from each other is over which patterns of action they claim are *favoured*. Kantianism has its own account of this, in terms of the value of good will and of respecting the humanity of rational agents. According to Kantianism, these considerations are such that KE patterns are the most favoured of all. Other theories of pattern-based reasons could give quite different

[9] C. M. Korsgaard, *Creating the Kingdom of Ends* (Cambridge, 1996), p. 150. Note that Korsgaard explains Kant's idea in terms of '[doing] my part'. I am grateful to Guy Fletcher for pointing out the relevance of this passage.

accounts. One salient distinction is between those theories that account for favouring entirely in terms of value or goodness (teleological theories) and those that appeal to some other consideration (non-teleological theories). I have defined Kantianism as a teleological theory (what favours the KE patterns is, exclusively, their value), though obviously it is controversial to interpret Kant's own views in this way.[10]

A second dimension of variation amongst theories of pattern-based reasons is over which patterns are *eligible*. 'Eligibility' here means that the pattern supports a reason (according to the theory). Most if not all theories of pattern-based reasons will place some constraints on eligibility, thereby allowing that some favoured patterns do not support reasons. Perhaps the least controversial constraint on eligibility is logical possibility: that is, a pattern's being logically impossible entails that it does not support any reason to realize any part of it. Kantianism endorses this constraint and adds some others. All the parts of any KE pattern must be compossible, for example. Moreover, the requirement that maxims can be willed as universal laws imposes a constraint that applies to the whole set of KE patterns. It requires that, if an action is part of a KE pattern for one agent in a certain circumstance, it must be part of some KE pattern for every other agent in circumstances that are, in other respects, morally the same.

Perhaps the most significant feature of Kantianism's account of eligibility, however, is that it claims that the eligibility of any KE pattern in no way depends on the likelihood of its realization as a whole. In particular, there is no connection between the eligibility of any KE pattern and the willingness of other rational agents to perform their parts of it. If there were such a connection, Kantianism would tie each agent's duties to the intentions of others. But Kantianism claims that in every circumstance, for every agent, at least one KE pattern is eligible. This is why every agent always has a reason to act as a member of the Kingdom of Ends.

Other theories of pattern-based reasons impose different constraints on eligibility. In particular, they may impose some form of *willingness requirement* – according to which a pattern is eligible only if (a sufficient proportion of) the other agents required to realize it are willing to realize their parts of it. This has the effect of drastically reducing the number of eligible patterns, of course, and also of mitigating worries about reckless unilateral action (though it is not the only way of doing this). Theories that accept some version of the willingness requirement are cooperative in the sense that (roughly speaking) they tell us to treat other agents differently according to whether they are cooperative

[10] See B. Herman, *The Practice of Moral Judgment* (Cambridge, Mass., 1993), ch. 10.

or not.[11] A different sort of eligibility constraint is the *first-person restriction*. According to this restriction, a pattern is eligible only if every part of it could be realized by the agent herself. Whereas the patterns of action we have so far discussed include the actions of other agents, the first-person restriction claims that the only eligible patterns are ones composed entirely of the agent's actions (such as plans of action).

The most restrictive constraint on eligibility is the *option restriction*. This claims that the only eligible 'patterns' of action are the agent's own options. There are significant difficulties in spelling out precisely what should count as the set of options for an agent in any given case.[12] But however that may be, it is clear that this is a very restrictive constraint. The patterns that it implies are eligible are themselves highly restricted – consisting either of actions the agent could immediately perform (on a very narrow understanding of 'option') or of relatively short sequences of actions that she could immediately begin (on a broader understanding of 'option'). At its most restrictive, the option restriction gives us a limiting case of the idea of pattern-based reasons, in which the action *A* for which there is a reason is identical to the favoured 'pattern' *P* (treating identity as a limiting case of parthood). Adding the option restriction thus generates a theory according to which we have pattern-based reasons only in the very restricted sense that we have reasons to perform some action *A* because this action itself is favoured. (These are what we might call 'act-based reasons': a special case of pattern-based reasons.) Though this is a limiting case, it is worth emphasizing here because it will enable us to treat act-utilitarianism as a theory of pattern-based reasons.

The third dimension of variation is over the *interaction* of the various reasons that exist, according to the theory. Many theories of pattern-based reasons might imply that an agent can have reasons associated with different sorts of pattern in a single case. For example, she might have one reason associated with a pattern involving many other agents, another associated with a first-person plan stretching over the rest of her life, and a third associated with her immediate options. Assuming that these reasons could conflict with each other, any theory of pattern-based reasons that aims at yielding deontic verdicts had better have something to say about which reason is strongest in cases

[11] See D. Regan, *Utilitarianism and Co-operation* (Oxford, 1980), p. 124.

[12] See L. Bergström, *The Alternatives and Consequences of Actions* (Stockholm, 1966); M. J. Zimmerman, *The Concept of Moral Obligation* (Cambridge, 1996); D. Portmore, 'What Are Our Options?', unpublished paper, <www.public.asu.edu/~dportmor/What%20Are%20Our%20Options.pdf> (2012).

of conflict between reasons.[13] It is important to see that this is a separate matter from the account of favouring. This is because the issue of how reasons interact is, in part, an issue about the relative importance of different reason-giving relations to favoured things: the parthood relation as compared with the causal relation, for example. Two theories could agree about favouring and eligibility yet disagree on the relative importance of parthood, say: one theory relegating pattern-based reasons to minor importance, while another treats them as always decisive.

For its part, Kantianism claims that the reasons associated with the KE patterns always defeat any other reasons an agent may have. This is (at least in large part) its account of interaction, and it explains why the agent ought always to act on these reasons, or always as a member of the Kingdom of Ends.

In sum, theories of pattern-based reasons differ from each other according to their accounts of favouring, of the eligibility of patterns and of the interaction of reasons. Kantianism is the result of distinctive accounts of these matters, according to which the KE patterns are most favoured, are always eligible and are associated with reasons that are always decisive. As I have tried briefly to indicate, many other views on each of these issues are possible.

III. ACT-UTILITARIANISM

It should by now be obvious that we can understand act-utilitarianism too as a theory of pattern-based reasons. The key to this is allowing *individual actions* to count as 'patterns of action'. Act-utilitarianism recognizes pattern-based reasons only involving patterns of this minimal sort ('act-based reasons').

Thus one element of act-utilitarianism is an account of eligibility that includes the option restriction. According to this restriction, only the agent's own options are eligible. Though she may have reasons, according to act-utilitarianism, to try to *cause* the realization of extended sequences of her own actions, or extended patterns of action performable by others, her reasons to do these things have nothing to do with any parthood relation between her options and these larger patterns.

Second, act-utilitarianism adopts a teleological and welfarist account of favouring, according to which the most favoured option is the one that

[13] I am assuming that what an agent ought to do is some function of her reasons. 'Interaction' is the function that takes us from reasons to oughts.

maximizes utility.[14] Third, act-utilitarianism adopts an account of the interaction of reasons according to which the strength of a reason to perform an option increases as the option is more favoured, and the agent ought to perform the option for which there is strongest reason.

The result of these claims is the theory according to which each agent ought, in every circumstance, to perform that option of hers that maximizes utility. Both Kantianism and act-utilitarianism are theories of pattern-based reasons. They differ from each other because they adopt different accounts of favouring, eligibility and interaction.

IV. THE PAY-OFF

So far I have claimed that Kantianism and act-utilitarianism can be understood as theories of pattern-based reasons, and I have explained the differences over favouring, eligibility and interaction that distinguish them when thus understood. But even if these claims are true, we might wonder about the point of understanding these theories in this way. Doing so requires introducing new terminology, which is costly. What is the pay-off? In this section I will cite five advantages of this way of understanding them.

First, this way of understanding them is charitable. Set aside, for the moment, the question about whether it is cumbersome or otherwise justified. One thing to be said for understanding act-utilitarianism and Kantianism as theories of pattern-based reasons is that doing so presents them sympathetically. Act-utilitarianism is presented as a theory that focuses on the good that can be achieved by individual actions, for example. This is just the way that defenders of act-utilitarianism typically present it.[15] Note that I do not claim that this way of understanding act-utilitarianism presents its appeal more clearly, or is an improvement on standard ways of presenting it. The current point is merely that the representation of it as a theory of pattern-based reasons preserves its appeal.

Similarly, the representation of Kantianism as a theory of pattern-based reasons preserves the appeal of at least some of Kant's ethical ideas. In particular, it offers one way of tying together three formulations of the categorical imperative. In realizing part of a KE pattern, an agent does the following. She acts on a maxim that could be willed as a universal law (no action is part of any KE pattern unless it is permitted by a maxim that passes this test). She acts as if she were

[14] We can set aside the issue of whether this should be understood as maximizing expected, or instead actual, utility.

[15] For example, J. J. C. Smart, 'An Outline of a System of Utilitarian Ethics', in J. J. C. Smart and B. Williams, *Utilitarianism: For and Against* (Cambridge, 1973), pp. 3–74.

legislating for the Kingdom of Ends (as before, her maxim could be willed as a universal law; in addition, her will is compatible with every other rational agent also acting as a legislator for the Kingdom of Ends). Third, her action respects the humanity of other rational agents (since her actions treat those others as potential co-realizers of a KE pattern, rather than as causal powers that could be recruited to promote her aims). On this interpretation Kant's ethics involves a common project, which has the special feature that each agent has sufficient reason to play her part in it whether or not others will play their parts.[16] This is an appealing way of understanding at least some of Kant's ideas.

A second advantage of this way of understanding these theories is that it emphasizes the common ground between them. What might otherwise seem like fundamentally different approaches to ethics instead appear, on this interpretation, to be the result of taking different views on the issues of favouring, eligibility and interaction. Seen this way, these differences are in-house disputes amongst those who share commitment to a common framework given by the idea of pattern-based reasons. Why might this help? It provides a way of framing disputes between them that should be acceptable to both sides. For example, Kantians could debate eligibility with act-utilitarians. Both could agree that reasons are associated with patterns of action. Each could try to advance reasons in favour of its account of which patterns support reasons. Though there is no guarantee that considerations could be found that would persuade the other side, this is an improvement over traditional encounters between Kantians and act-utilitarians, which are often marked by mutual incomprehension. When both theories are understood as theories of pattern-based reasons, they appear to differ in reasonable ways over common issues.

Third, understanding Kantianism and act-utilitarianism as theories of pattern-based reasons helps to explain some important differences between them. One such difference is the prominence within Kantianism, but not act-utilitarianism, of the idea of imperfect duties. Why is this idea prominent in one theory but not the other? The answer has to do with the theories' accounts of eligibility and favouring. Kantianism treats extended patterns of action as eligible, while act-utilitarianism does not. Moreover, its account of favouring treats several patterns as equally good (for example, any pattern of action through which the agent sufficiently develops her talents is equally

[16] Kant writes: 'a kingdom of ends would actually come into existence through maxims whose rule the categorical imperative prescribes to all rational beings *if they were universally followed*' (*Groundwork* 4: 438, in M. J. Gregor (trans. and ed.), *Immanuel Kant: Practical Philosophy* (Cambridge, 1996), p. 87, emphasis in the original).

good in that respect). The result is that Kantianism applies duties to extended patterns (forbidding some lives, for example) in a way that does not mandate specific actions on specific occasions (since the agent is permitted, on any occasion, to realize part of any of the favoured patterns). This combination – certain sequences are prohibited, but no constituent member of a sequence is, just for that reason, prohibited – is described by the idea of an imperfect duty. There is no particular day on which I am required to read the *Methods of Ethics*, just so long as I do not always forgo reading it.

Act-utilitarianism has much less room for indifference between 'patterns' (that is, on its view, options). However, it does recognize the possibility of tied optimal options. This is the analogue in act-utilitarianism of the idea of imperfect duties. What it says about these cases is similar to what Kantianism says about imperfect duties: there is a duty to realize one of the favoured patterns. As with the case of developing one's talents, the case of tied optimal options leaves the agent with moral freedom to choose between permitted actions. But in the case of tied optimal options under act-utilitarianism, there is not the further complexity that there are several permitted options that could be taken in order to realize part of a permitted pattern – since each permitted pattern just is a permitted option. So the idea of tied optimal options is the attenuated form that the idea of imperfect duties takes within act-utilitarianism.

Another difference between Kantianism and act-utilitarianism is even more striking and important. This concerns the moral significance each theory assigns to the *response* of others to the agent's actions. Act-utilitarianism is interested in the way in which others *would* respond to the agent's actions. The utility of each of the agent's options depends on the response it would elicit from those others who stand to affect the success of her actions. In this respect, the behaviour of others in the agent's environment has the same status, for act-utilitarianism, as does the behaviour of any other practically relevant part of her environment (including inanimate parts). This focus on *what would happen* gives act-utilitarianism its characteristically pragmatic nature. It makes act-utilitarianism's practical implications generally very sensitive to changes in others' intentions. But now contrast this with Kantianism. The way that others would respond to the agent's options does not affect the moral worth of those options, according to Kantianism. This is most dramatically illustrated in the case of the murderer at the door. Consider variants of this case, in which the response of the person at the door varies: from showering the recipient with riches and love, through giving him some minor reward, through inflicting some minor harm, to murder. Travelling from one end of this spectrum to the other – which is to say, varying how the other person *would respond* to being

told the truth – makes no difference to the duty not to lie, according to Kantianism. Instead, what matters is only how the other person *could respond* – which on Kantian assumptions remains constant, and always includes acting as a member of the Kingdom of Ends.

This is a tremendously important difference between these theories. Tying duty to how others could respond instead of how they would respond is the cause of many of the differences in practical implications between Kantianism and act-utilitarianism. Moreover, it explains some of the criticisms each view has received. When Bernard Williams criticizes act-utilitarianism for undermining the agent's integrity, he points to cases in which the agent's duty depends on how other agents (not inanimate parts of the environment) would respond. Contrary to his diagnosis, the fundamental issue is not the way in which act-utilitarianism treats the relationship between the agent and her own actions, but instead the way it treats the relationship between her duty and the response of other agents. In tying her duty to their response, it makes her duty dependent on their behaviour, which encourages her to take an instrumental attitude to her own actions. Intuitively, adjusting one's actions because of the anticipated behaviour of an inanimate part of one's environment is less of a threat to one's integrity than doing the same in response to another agent's change of intentions.[17] In contrast, Kantianism insulates duty from facts about how others would respond. This gives each agent a kind of moral independence that protects, to the highest degree possible, what Williams calls 'integrity'. But the downside is that it also involves setting aside facts about the consequences of the agent's options that, intuitively, we think it wrong to set aside.

On standard interpretations of Kantianism and act-utilitarianism, their differential focus on what others could or would do is a puzzling and unexplained difference between them. But it is elegantly explained if we understand them as theories of pattern-based reasons. Any theory of pattern-based reasons distinguishes between two categories of behaviour. First, there is the behaviour of agents ('insiders') whose action is included in an eligible pattern. Second, there is the behaviour of agents ('outsiders') whose action is not included in an eligible pattern. So far as the behaviour of insiders goes, what matters is what the agent concerned *could* do. This follows from the constraint on eligibility, which is accepted by most theories of pattern-based reasons including Kantianism, that no pattern of action is eligible unless all of its component parts could be performed. Whether or not the agent

[17] B. Williams, 'A Critique of Utilitarianism', *Utilitarianism: For and Against*, pp. 75–150, esp. 98–116. This may be why Williams's examples all involve problems caused by other agents.

would perform it is a different matter, and not at issue.[18] However, this is in contrast to what matters about the behaviour of outsiders. Their behaviour could affect the consequences, and so the value, of the whole pattern. So if their behaviour matters at all, what will matter is how they *would* respond. Thus, in general, theories of pattern-based reasons will tie reasons and duty to what insiders could do and what outsiders would do. Now apply this general rule to Kantianism and act-utilitarianism. According to Kantianism, *all* rational agents are insiders: none are outsiders. So, for all rational agents, what matters according to Kantianism is how the agent could respond. But according to act-utilitarianism, all other agents are outsiders: *none* (except the primary agent's present self) are insiders. So, what matters according to act-utilitarianism is how they would respond to each of the primary agent's options. (Its concern with these options is the limit of its concern with what any agent could do.) This is why Kantianism and act-utilitarianism differ over the response of other agents, with all of the consequent differences for their practical implications. This is an insight into the different characteristics of Kantianism and act-utilitarianism which is accessible to us only because we can understand both of them as theories of pattern-based reasons.

A fourth advantage of understanding Kantianism and act-utilitarianism in this way is that it explains their relationship to some other important theories. Of course we can see the relationship between act-utilitarianism and act-consequentialism: act-utilitarianism is the result of combining act-consequentialism with a particular theory of favouring. Less obviously, we can see the relationship between act-utilitarianism and some of its other cousins. Standard forms of act-utilitarianism incorporate the option restriction in their accounts of eligibility, which has the result that they treat the primary agent's own future actions as the actions of an outsider. In contrast, *possibilist* theories include the first-person restriction but not the option restriction. The result is that all and only the actions performable by the primary agent are treated as the actions of an insider. This difference is salient when the primary agent could but would not make some beneficial response to a present option.[19] As mentioned above, *cooperative* forms of utilitarianism adopt a different kind of constraint on eligibility. They adopt some form of the willingness requirement, according to which a pattern is eligible only if (a sufficient proportion of) the other agents required to realize it are willing to realize their

[18] Theories that include a willingness requirement are a complicated partial exception.

[19] Zimmerman, *Moral Obligation*, ch. 6; F. Jackson, 'Group Morality', *Metaphysics and Morality: Essays in Honour of J. J. C. Smart*, ed. P. Pettit, R. Sylvan and J. Norman (Oxford, 1987), pp. 91–110.

parts of it. These theories set the insider/outsider boundary according to facts about willingness.[20]

Possibilism and cooperative-utilitarianism differ from act-utilitarianism only because of their accounts of eligibility. We might assume that the same is true of rule-utilitarianism. It presumably adopts an account of eligibility that denies the option restriction, the first-person restriction and the willingness requirement (since it considers the consequences of agents' compliance with sets of rules even when those agents are unwilling to comply). But it might be that rule-utilitarianism differs in another important way too. Leading forms of rule-consequentialism tend to include a disaster-prevention rule, which permits breaking any other rule if necessary to avoid disaster.[21] Supposing that rule-utilitarianism includes this feature, it might be best understood as a pluralist theory of pattern-based reasons. On this interpretation, most rule-utilitarian reasons for action are associated with extended patterns of action consisting of near-universal compliance with rules. But the reason to do what is necessary to avert disaster is associated with a minimal pattern: the disaster-averting option itself. (The reason to take some action that would avert disaster is that *this token action* would have much better consequences than the alternatives; it is not a reason to play one's part in a larger pattern of disaster-aversion.) This illustrates an important theoretical possibility. Theories of pattern-based reasons can be pluralist in the sense of claiming that there are reasons in a single case associated with patterns of quite different extents. Such theories incur the burden of explaining how these different reasons interact. Disaster-averting rule-utilitarianism presumably incorporates an account of interaction roughly like this: reasons associated with extensive patterns typically defeat reasons associated with minimal patterns; however, when disaster is at stake they are defeated by the reasons associated with minimal patterns.[22]

One might wonder whether understanding these relationships with other views really improves our understanding of Kantianism and act-utilitarianism themselves. In reply, note that it is widely accepted that understanding act-utilitarianism as a form of act-consequentialism is worthwhile because this reveals the relationship between act-utilitarianism and closely related theories that employ different theories of value. The claims I have just made have the same form. It is worthwhile understanding Kantianism and

[20] Regan, *Utilitarianism*, p. 124; E. F. McClennen, *Rationality and Dynamic Choice* (Cambridge, 1990).

[21] For example, see B. Hooker, *Ideal Code, Real World* (Oxford, 2000), pp. 98–9.

[22] Woodard, 'Rule Consequentialism', pp. 257–8.

act-utilitarianism as theories of pattern-based reasons because this reveals the relationships between these theories and closely related theories that employ different theories of favouring, eligibility or interaction (or some combination of these). In both cases, the claim is that we get some insight into theory *X* by seeing how it is related to some other theories *Y* and *Z*.[23]

The fifth advantage of understanding Kantianism and act-utilitarianism as theories of pattern-based reasons is that this enables us to see how to generate novel ethical theories worthy of consideration. The method is as follows. Take any well-known theory of pattern-based reasons, and modify its claims in one or more of the three dimensions (of favouring, eligibility and interaction) in ways that are plausible but that do not merely reproduce views that are already well known. The result will be a close relative of well-known views, which is likely to have interestingly different properties. We do not have any good reason to assume that the theories that are already well known are the only ones worth studying.

In sum, we have the following reasons to understand Kantianism and act-utilitarianism as theories of pattern-based reasons. Understanding them in this way is charitable. Second, it emphasizes what they share in common. Third, it explains some important differences between them, such as the prominence in Kantianism but not act-utilitarianism of the idea of imperfect duties, or (very significantly) their different treatment of the behaviour of other agents. Fourth, it enables us to understand their relationship to some other important theories, such as possibilism, cooperative-utilitarianism and rule-utilitarianism. Finally, it gives us a method for generating new views worthy of study.

V. TWO OTHER PROPOSALS

It may be helpful to explain how my proposal is related to some other recent proposals about the structure of Kantianism and act-utilitarianism. I will briefly discuss two such proposals: the consequentializing idea and Michael Ridge's proposal about 'Consequentialist Kantianism'.

According to the consequentializing idea, every theory of normative ethics can be 'consequentialized': that is, accurately represented as a version of act-consequentialism.[24] To consequentialize a theory, one must devise a theory of value that, when combined with act-consequentialism, yields the target theory's deontic verdicts in all cases. If the consequentializing idea is correct, one could regard

[23] I am grateful to an anonymous referee for pressing me on this point.
[24] Portmore, 'Consequentializing', pp. 39–40.

act-consequentialism as the common structure of Kantianism and act-utilitarianism (and of every other theory of normative ethics).

Consider roughly how this would work for one element of Kantianism: the exceptionless prohibition of lying. For a consequentialized version of Kantianism to yield a prohibition of lying, the theory of value must assign great disvalue to the act of lying. This would help to explain why lying is wrong even when it could bring rewards, for example. But of course the Kantian prohibition requires more than this. It requires that lying is prohibited in every case – including cases in which my lying now could prevent you telling several lies later, and cases in which my lying now could prevent me telling several lies later. Thus the theory of value also has to attribute much greater disvalue, so far as my duties now are concerned, to my lying now than to your lying later or my lying later. In other words, the theory of value must distinguish acts according to when and by whom they are performed, and assign greatly different values to such acts even when they are, in other respects, morally very similar. Such a theory would make claims about the value of actions that, taken by themselves, would be very hard to believe.[25]

None of this is to dispute whether the consequentializing idea is correct. The idea claims only that it is possible to represent every view accurately as a version of act-consequentialism, and it treats deontic equivalence as sufficient for 'accurate representation'. Understood in those terms, the consequentializing idea may be correct; and if it is, that may have important consequences for ethical theorizing.[26] However, it is important to note that my proposal has had a more ambitious aim. I claimed that Kantianism (as I stipulated it in section I) is a *charitable* interpretation of some of Kant's ideas. It could not be claimed that a consequentialized version of Kant's ethics is a charitable interpretation of it. As we have seen, the theory of value it would assign to Kantian ethics would be very hard to believe. The credibility of a theory does not depend only on the deontic verdicts it yields. This is why maintaining deontic equivalence is not sufficient to maintain credibility, and why consequentializing a theory can amount to uncharitable interpretation of it.

The consequentializing idea tries to fit other theories into the structure of act-consequentialism. As I have explained, act-consequentialism can be understood as a theory of pattern-based reasons (albeit one that recognizes only minimal patterns as eligible). Hence, the consequentializing idea is a specific instance of my proposal.

[25] Ridge makes a similar point. See Ridge, 'Consequentialist Kantianism', pp. 425–6.
[26] Portmore, 'Consequentializing', pp. 41–3.

But it proceeds by holding two of the three dimensions of variation I identified (eligibility and interaction) fixed, and allowing variation only in the third (favouring). Because it does not make use of these other dimensions of variation when representing other theories, it is easy to understand why it can deform them. By contrast, my proposal has more resources to represent other theories in a way that interprets them charitably.

Consider this point in relation to the explanation of exceptionless duties, such as the duty not to lie. The challenge for any teleological construal of Kantian ideas is to explain how the badness that explains why the agent should not lie fails to support a reason for her to minimize lying. The consequentializing idea attempts to meet this challenge by distinguishing between the badness of her lying now and the badness of others lying or of her lying later. But there is no need to make these manoeuvres within value theory if we have recourse to other structural resources. The account of Kantianism that I have given meets the same challenge in the following way. It distinguishes between reasons associated with KE patterns (which are reasons for each agent not to lie on any particular occasion, since lying is incompatible with all such patterns) and reasons associated with minimal patterns (which may be reasons to bring it about that there is less lying in future), and it claims further that the reasons associated with KE patterns always defeat any reasons associated with minimal patterns. There is no need, on this account, to postulate different kinds of badness according to who is lying or when it occurs. Instead, there are different morally significant relations (parthood, cause) that can hold between an agent's options and the badness of lying.

Recently, Michael Ridge has made an ingenious proposal that is roughly the opposite of the strategy pursued by the consequentializing idea. It allows us to treat Kantian ethics as a consequentialist theory whilst combining it with a teleological and agent-neutral theory of value.[27] On this interpretation, Kantianism has a theory of value that has much in common with the theory employed by act-utilitarianism. Instead, the distinctively Kantian ideas are modelled using a non-standard theory of rightness. According to this 'mini-min' theory, each agent ought to minimize the chance of the worst possible outcome. The worst possible outcome (according to the theory of value postulated for Kantianism by Ridge) is one in which every agent has a bad will. What's more, each agent has control only over her own will. She cannot ensure that any other agent does not have a bad will. Thus, the mini-min theory

[27] Ridge, 'Consequentialist Kantianism'. By 'teleological', Ridge means a conception of value according to which it is to be promoted (p. 424). Note that this is more restrictive than my usage of 'teleological' in sect. II, above.

Table 1. *Ridge's consequentialist Kantianism as a theory of pattern-based reasons*

Favouring	The best pattern is the one that minimizes the chance of every agent having a bad will
Eligibility	Only minimal patterns (options) are eligible
Interaction	The reason associated with the best pattern defeats all other reasons

of rightness, together with the postulated theory of value, implies that each agent ought to ensure that she does not have a bad will herself, since this makes the worst possible outcome impossible and nothing else she could do would achieve as much. As Ridge helpfully puts it, the agent is especially concerned with the goodness of her own will not because it is more valuable, but only because (on Kantian assumptions about free will) she cannot completely control any other will.[28]

As Ridge presents his suggestion, what 'consequentialist Kantianism' shares in common with other act-consequentialist theories is that it focuses on what each action taken by itself achieves, and it employs a teleological and agent-neutral theory of value.[29] The mini-min theory of rightness is presented as being altogether alien to consequentialism. However, as Table 1 indicates we can in fact assimilate the mini-min theory of rightness within the structure I have described, and which I have claimed is shared by act-utilitarianism.

Though Ridge's consequentialist Kantianism fits within the structure I have described, it does not take full advantage of the resources of that structure. It offers an ingenious way of modelling Kantian endorsement of exceptionless prohibitions. But it cannot properly model another important idea in Kantian ethics: the idea of imperfect duties. It *can* appeal to an account of what is necessary for a good will that stipulates that complying with the imperfect duties is necessary, but it does not seem able to explain how that requirement is associated with any reason to engage in self-improvement or acts of beneficence on any particular occasion. The interpretation of Kantianism that I have offered has an advantage in this respect, since it associates reasons with the opportunity to realize parts of favoured patterns. This is due to its use of a less restrictive account of eligibility.

[28] Ridge, 'Consequentialist Kantianism', p. 435.

[29] Note that Ridge appears to be restricting 'consequentialism' to theories with agent-neutral theories of value. In contrast, the consequentializing idea relies on the supposition that consequentialism is compatible with agent-relative theories of value.

VI. CONCLUSION

Kantianism shares a structure with act-utilitarianism in the sense that both are theories of pattern-based reasons. They differ from each other according to the accounts of favouring, eligibility and interaction of reasons that they each employ. This structure gives us a way of framing the differences between them that is charitable to each theory, that highlights what they share in common and gives us new insights into their differences, that shows how they are related to other important theories, and that gives us a procedure for generating new theories worth studying.

My proposal applies specifically to 'Kantianism', the view I stipulated in section I. I have not claimed that Kantianism is Kant's own view, or that it captures the richness of his ideas. However, the interest of my proposal depends in part on resemblance between Kantianism as I have defined it and Kant's own views. If Kantianism were to have little to do with Kant's views, it would be less interesting to find that it shares a structure with act-utilitarianism.[30]

christopher.woodard@nottingham.ac.uk

[30] For helpful comments and discussion I am very grateful to Andrew Fisher, Guy Fletcher, Carrie Ichikawa Jenkins, Daniel Nolan, Douglas Portmore, Neil Sinclair, members of the audience at the Rocky Mountain Ethics Congress, University of Colorado, Boulder, on 7 August 2009 (especially Julia Driver, Shelly Kagan and Arthur Ward), members of the audience at the University of Manchester research seminar on 12 October 2011, and anonymous referees for this journal.