

## CHAPTER IV

### MORAL UNIVERSALISATION : KANT AND HARE

Almost everyone now agrees that moral judgements must, in some sense, be universalisable. And this agreement is typically taken to be an agreement with a central contribution of Kantian ethics. Much that is currently said about moral universalisation, however, would have been strongly rejected by Kant. And so the label 'universalist' does not of itself aid our understanding of Kant's position at all and can easily interfere with such understanding.

R.M. Hare, for example, offers a theory of moral universalisation that is often regarded as Kantian in character. According to Hare, the fundamental principle of Morality is the principle that whatever rule any agent applies to other persons he must also apply, or be willing to have applied, to himself, and conversely. This is, according to Hare, simply the demand that moral judgments must, as a requirement of rationality, be universalisable. This theory is quite clearly formal (i.e. any end or purpose can be a part of morality so long as the agent wills it for everyone including himself), and it does appear initially plausible to say that it enshrines a kind of Kantianism. But this is a misleading impression.

First of all, it is important to see how very strange Hare's view really is. For it is really quite radically subjective

in character. For Hare, the universalisability of a judgment depends solely upon what the agent is willing to accept. He is not claiming that a universalisable moral judgment is one which in fact could obtain as a universal practice. Rather it is one which the agent does will to accept as a universal practice. No matter how evil or unworkable the state of affairs, if the agent is willing that he and everyone else labour under it, then the judgment is moral. The universalisability of a judgment, then, is not determined by any objective state of affairs in the agent's environment, but only by what the agent is or is not willing to put up with. 'The test of the agent's wished for himself qua recipient would justify rules which impose on their recipients unjust or immoral hardships, including racial discrimination and even genocide. If Werner is willing to be exterminated if it is discovered that he is a Jew, then Werner's prescriptive judgment 'Exterminate the Jews' counts as moral. Morality, on this view, becomes essentially a private rather than a public enterprise. As H. L. A. Hart remarks:

To characterise morality ... as primarily a matter of the application to conduct of those ultimate principles which the individual accepts or to which he commits himself for the conduct of his life seems to me an excessively Protestant approach. Important as this aspect or kind of moral judgment is, we need to understand it as a development from the primary phenomenon of the morality of a social group ('Legal and Moral Obligation' in Essays in Moral Philosophy).

Now it must be admitted that Kant's statement of the Categorical Imperative would, if viewed uncritically, incline one to interpret Kant as holding a similar position:

Act only according to that maxim by which you can at the same time will that it should become a universal law (Foundations, 421 : Beck, 39).

This certainly sounds as though Kant, like Hare, were saying that you act on moral (as opposed to private) grounds so long as you are simply willing that your maxims be universal. However, we should be suspicious of pinning such a view on Kant, for we must remember that his exercise in pure moral philosophy is to discover the nature of fully rational action. And the imposition of genocide by a fanatic hardly seems to correspond with such an Ideal paradigm of rationality - no matter how much the fanatic is willing to suffer the awful consequences of what he does. And indeed, if we proceed a few paragraphs further in the Foundations, we find that Kant's view is considerably more complex than Hare's:

We must be able to will that a maxim of our action become a universal law; this is the canon of the moral estimation of our action generally. Some actions are of such a nature that their maxim cannot even be thought as a universal law of nature

without contradiction .... In others this internal impossibility is not found though it is still impossible to will that their maxim should be raised to the universality of a Law of nature, because such a will would contradict itself (Foundations, 424; Beck, 41-2).

So Kant is really operating here with two principles of morality. The first, used to derive perfect duties when later employed as a criterion, is that you should act on no maxim which is incapable of being a universal practice. The second, later used as a criterion for imperfect duties, is that certain maxims, even if capable of being universal practices, cannot count as moral if the agent cannot consistently will that they be universal practices. And neither of these principles is subjective. The first speaks, not just of what can be willed (for, as the existence of irrational people surely shows, anything can be willed), but what can be consistently willed. And, as I shall argue later, by 'consistently willable, Kant does not mean merely consistency with whatever contingent desires I, by some personal quirk, just happen to have. Rather he means consistency with the desires that all men necessarily have. This desires he calls essential ends of humanity, and they play a crucial role in Kant's material metaphysics of morals.

Kant's views on moral universalisation, then, are not formal in the sense that Hare's are. But what are his views

then? To see this, we shall need to examine the three pure formulations of the Categorical Imperative. These are all recognisable by the fact that they make no reference to nature or to humanity, but are stated as characteristic of rationality as such. They may all be expressed in the indicative mood, and they should be so expressed to avoid confusion between pure and applied moral philosophy. For my working formulations, I shall use the following:

1. X is a fully rational being if and only if X acts on maxims that are universalisable.
2. X is a fully rational being if and only if X treats rationality, whether in its own being or in that of another, always as an end never as a means only.
3. X is fully rational being if and only if X acts as though he were a law-making member in a universal kingdom of ends.

What I wish to argue in the remainder of this chapter is that formulation (2) is the fundamental formulation of the Categorical Imperative. It, and the doctrine of 'ends in themselves' upon which it rests, are absolutely essential to understanding what Kant means by moral universalisation.

Now the deduction of formulation (1) from the concept of a rational being is easily presented. Kant does not actually give us the argument in the Foundations (except as an enthymeme),

but by adding a premise (b) from the Critique of Pure Reason, we can get the following:

- (a) X is a fully rational being if and only if X acts only according to a conception of law (at Foundations, 412; Beck, 28).
- (b) X is a law only if X is universal (at Critique of Pure Reason, A 2; Kemp Smith, 42).

Therefore : X is a fully rational being if and only if X acts according to a conception of universal law.

We seem to have, then, as a fundamental principle of a fully rational being, that it acts only on a conception of universal laws. However, having this principle is not in itself any great asset in our inquiry into pure moral philosophy. For, as it stands, the notion of 'universal law' lacks specification and cannot be brought to bear upon specifically moral issues. For surely we would not want to claim, for example that there is any moral significance in acting on a conception of the law of universal gravitation - though such action certainly satisfies a bare criterion of universalisability.

This difficulty, however, is more apparent than real : For we must remember that we are talking about the universalisability of maxims - principles of human action. But behaving in accordance with the law of universal gravitation has no maxim because such behaviour is not, in any meaningful sense, an

action at all. For it is not within my power to refrain from behaving in accordance with this law. Though my jumping off a building might very well be an action of mine, we should hardly want to say that my falling and hitting the ground was also an action on my part. It is not something I did, but something which happened to me. Thus in spelling out universalisability in any ethically relevant sense (in a sense characterising intentional human actions), we must remember that what must be universalisable is not brute bodily behaviour as such, but the maxims of actions. Universalisability thus cannot be spelled out as simply consistency with actual laws of nature. I act according to these, surely, but not on a conception of them.

Neither can 'universalisable' mean 'logical consistency' in any formal sense. We can know whether a given statement is formally consistent without knowing anything about the content of the statement. This is, after all, the value of formal procedures. They test the logical consistency of statements once the actual terms or content of these statements have been replaced by variables. For example, we know that the statement 'All grinchies are greeps and there is one grinch which is not a greep' is formally inconsistent without knowing anything at all about the meaning and content of the actual statement. For the general scheme 'All Ps are Qs and there is a P which is not a Q' is formally inconsistent, i.e. it is reducible to a contradiction of the form 'P and not P'. But the universalisability

of maxims does indeed depend upon the content of those maxims. Universalisability is supposed to be a rational criterion of conduct, not of statements, and must take account of the ends that the agent is pursuing in action. The maxim 'I shall make a false promise' is, according to Kant, not universalisable. But it obviously is not a formal contradiction, nor can it be reduced to one. What is not universalisable in this maxim is not its logical form, but is rather the end or purpose that the agent is seeking to bring about.

My reason for discussing these points has been to show that universalisability cannot in itself stand as a sufficient condition for rational morality. For, if taken as sufficient, the sphere of morally permissible actions will include those that merely accord with natural causal laws (e.g. 'survival of the fittest') and those which can be given a non-contradictory description (e.g. 'kill the Jews'). Thus the notion of universalisability must, if it is to be of any help at all in determining a characterisation of the moral point of view, be spelled out so as to be explicitly relevant to maxims-principles for the realisation of certain ends, purposes of states of affairs. All maxims are of the form 'To bring about so and so under certain conditions', and we need some principle to tell us the difference between what is rational to bring about and what is irrational to bring about. Moral actions must be universalisable or rationally consistent - but consistent with what?

Here real difficulties begin to present themselves. We must remember that Kant is trying to present his characterisation of the supreme principle of morality as a part of pure moral philosophy. Thus he cannot spell out universalisability in terms of any particular empirical ends or purposes that rational beings just happen to have, e.g. happiness. Rather he seeks a characterisation of the actions of all rational beings, and there is no reason to suppose that all rational beings pursue the same material ends. And even if they did, this would still be a contingent matter and could not form the basis for the a-priori principles of pure moral philosophy. But (and this is a very important 'but') if there were an end actually set by reason itself, then the case would be quite different:

It is necessary law for all rational beings that they should always judge their actions by such maxims as they themselves could will to serve as universal laws? if it is such a law, it must be connected (wholly a priori) with the concept of the will of a rational being as such. But, in order to discover this connection we must, however reluctantly, take a step into metaphysics.

Material ends ... are without exception only relative, for only their relation to a particularly constituted faculty of desire in the subject gives them their worth. And this worth

cannot, therefore, afford any universal principles for all rational beings.

But suppose that there were something the existence of which in itself had absolute worth, something which as an end in itself, could be a ground of definite laws. In it and only in it could lie the ground of a possible categorical imperative, i.e. of a practical law .