

WHAT IS AT STAKE. At the end of Section II (see below) Kant announces that the formulae of the categorical imperative and the theory of duty presented in the first two sections are incomplete in one key regard. 445

Kant interprets his procedure in these sections as the development of a generally accepted (common) morality and claims that the autonomy of will is the outcome of that development. In other words, the only possible morality is the one given by the categorical imperative. But we have not shown that *any* morality is real, or as he puts it, that it is not a “phantom of the brain”. This will be the subject of Section III.

Now, as we interpreted the discussion of FH, there was already a proof that categorical imperative applies to the human will, that it does have the power of command. Kant’s current concern, as I understand it, may loosely be paraphrased thus:

(6-1) Even if the will is governed by the categorical imperative, and even if morality (explicated with the concept of duty) is constituted by commitment to the categorical imperative, is it *actually* so that human beings do have the will, and that, therefore, they are bound by the categorical imperative?

So I think we can profitably interpret the task of Section III as engaging with a moral sceptic. He is a figure familiar from Plato, Hobbes, and Hume. He does not quarrel with particular formulae of duties. But he doubts whether any human action could satisfy or violate these duties. Compare him to an ‘inquisitive atheist’ who is willing to debate this or that religious duty, but at the same time denies the existence of God. Duties, according to him, are conditional on the reality of God (morality):

(6-2) There are duties if, but only if, there is God (morality). Since, however, there is actually no God (morality), there are no duties either.

Especially in Plato and Hobbes, having dismissed morality, the sceptic then proceeds to locate the real motivators of agency, which he usually finds in ‘happiness’ and self-preservation. It is unclear whether Kant has anything to say about this further move (most commentators claim he does not, and has never meant to, but I am not so sure).

Accordingly, section III can be split into two unequal parts. In the first half spanning the first three paragraphs Kant argues, roughly, that freedom is necessarily correlated with morality. In the second half spanning most of the remainder Kant argues that freedom can be ascribed to human beings.

FREEDOM AND ALIEN CAUSES. Right at the start the will is defined as a “kind of causality” possessed by rational beings. This definition is problematic, in the first place, because so far we thought of the will in teleological terms. We thought that willing is characterised by purpose: for example, ‘My will is to quit smoking.’ That is why we were easily able to talk about ‘good will’, meaning here the will that has a good purpose. If the will is conceived causally, then it is not easy to talk about ‘good causes’ or ‘good causality’. Other interpretations of Kant identify the will with character traits. This is consistent with our thinking of imperatives in Handout 3 in terms of justification. What ultimately justifies (explains) the action is the agent’s deeper commitments, his character. But now, can character traits be a form of causality? It may then be best to not try explicating causality or willing in more primitive terms and hope that the theory makes sense even without such explanations. 446

At all events, Kant’s intended contrast is between the power of the will and the power of desire. A parallel contrast is between freedom as an effective cause and causality according to nature (natural necessity). Observe Kant’s usage here (as elsewhere in the book): desires are characterised as “alien causes”. One might ask why they are ‘alien’. Why are my desires any less mine than my will? 446

Question 1. Give an example illustrating this contrast.

Remark 2. On the subject of alien desires and its relation to the will see Harry Frankfurt, ‘Freedom of the will and the concept of a person.’

THE NATURE OF FREEDOM. To say that to be free is to be not determined by natural causes tells us what freedom is not. But we want to know, positively, what freedom is. Kant interprets this question as the question of what laws govern free activities. This sounds like a contradiction: could there be laws of freedom? Kant claims there should be, since freedom is still a form of causality, and any causal behaviour should follow laws. 446

THE IDENTITY THESIS. Recall that, according to FA, to act morally is to act on the autonomous determination of your will. This implies that acting morally involves the ability to resist, with your will, the influence of desires. So it follows already from the FA that freedom would be necessary for morality. Now Kant wishes to argue that it is also sufficient. Therefore, a free will and a will acting under moral laws are identical. We have:

Identity thesis Free will is a necessary and sufficient condition for autonomous action and a necessary and sufficient condition for moral action: Freedom \Leftrightarrow autonomy \Leftrightarrow morality.

The argument for the Identity thesis is not difficult to make sense of, but rather more difficult to defend. It may be presented as follows:

446–47

- (1) A free will must be law obedient.
- (2) However, a free will cannot obey natural laws.
- (3) Therefore, it must obey the laws generated by itself (i.e. it must be autonomous).
- (4) Its autonomous laws are regulated by the principle of autonomy.
- (5) But the principle of autonomy is a moral principle (i.e. it is expressible in the FA equivalent to the FUL).
- (6) Therefore, a free will is governed by a moral principle.

The problem with this argument is evidently in the transition from the concepts of causality and determinism to the concepts of morality. In particular, self-determination can be understood in accordance with FA (putatively equivalent to FUL), or it can be understood as action under self-generated laws. There is no guarantee that those laws would coincide (let alone, necessarily coincide) with the laws prescribed by FA.

IN DEFENCE OF THE IDENTITY THESIS. The main problem with the thesis, as we just said, is the apparent equivocation Kant allows in the meaning of freedom and causality. How can we deal with this problem?

Free agents are able to resist their impulses, yet they can still *act on* their impulses if they ‘process’ them to include in their maxims. This processing capability requires freedom—simply because the agent is not controlled by his impulses. Rather, it is *he* who controls them.

It is instructive to compare such a free agent to a strategic hedonist. The latter is able to resist certain impulses for short-term gains in order to achieve long-term gains. But why does he frame his long-term gains in terms of pleasure satisfaction? No doubt he does this, we say, because ultimately he succumbs to his impulses however cleverly arranged.

This leaves open the further question why autonomy is determined exclusively by the moral law expressed in the categorical imperative. That is, why does the principle adopted by a rational agent coincide (by necessity) with the categorical imperative? It is perhaps easy to see why the moral law would be a sufficient condition for such a principle. If my maxim is right for all rational agents, then it is right for me, too.

But why would the moral law be a necessary condition for the principles adopted by the rational agent? Presumably this would exclude many innocent principles of skills (i.e. maxims governed by hypothetical imperatives). Well, in the light of our earlier discussion and the idea of categorical imperatives as ‘tests’, we can now say that the *non-violation* of the moral law should be a necessary condition for the principles adopted by the rational agent.

PROBLEMS. There are two possible problems with Kant’s account I want to mention here. (1) Kant has identified non-moral actions and immoral ones with those determined by ‘alien causes’. If so, they cannot really be attributed (or ‘imputed’) to the agent, with the result that the agent should not be blamed for them. (2) On the other hand, moral actions are identified with those determined by the free will. But the free will is, again, a kind of causality. So even moral actions cannot be attributed to the agent, with the result that the agents should not be praised for them. (This has to be elaborated further...)

THE REALITY OF FREEDOM. Here we are concentrating on Kant’s assertion that freedom should necessarily be ascribed to rational beings. The argument is very quick: it is easier to say what Kant does *not* claim than what he actually claims. A sketch of the argument is as follows:

448

- (1) We necessarily ascribe to rational creatures with a will the idea of freedom under which they act.
- (2) To act under the idea of freedom is to be free practically.
- (3) To be free practically is to be subject to the same laws as one would have been subject to if the will were shown to be free on theoretical grounds.
- (4) The free will is necessarily subject to the moral law (the Identity thesis).
- (5) Since rational creatures have free wills, they are subject to the moral law.

even though it decides nothing there it still preserves the indeterminate idea (of a will good in itself) unfalsified, for closer determination.

For the rest, I believe I may be excused from a lengthy refutation of all these doctrines.^k That is so easy, and is presumably so well seen even by those whose office requires them to declare themselves for one of these theories (because their hearers would not tolerate suspension of judgment), that it would be merely superfluous labor. But what interests us more here is to know that all these principles set up nothing other than heteronomy of the will as the first ground of morality, and just because of this they must necessarily fail in their end.

4:444

Wherever an object of the will has to be laid down as the basis for prescribing the rule that determines the will, there the rule is none other than heteronomy; the imperative is conditional, namely: *if or because* one wills this object, one ought to act in such or such a way; hence it can never command morally, that is, categorically. Whether the object determines the will by means of inclination, as in the principle of one's own happiness, or by means of reason directed to objects of our possible volition in general, as in the principle of perfection, the will never determines itself *immediately*, just by the representation of an action, but only by means of an incentive that the anticipated effect of the action has upon the will: *I ought to do something on this account, that I will something else*, and here yet another law must be put as a basis in me, the subject, in accordance with which I necessarily will this something else, which law in turn needs an imperative that would limit this maxim. For, because the impulse that the representation of an object possible through our powers is to exert on the will of the subject in accordance with his natural constitution belongs to the nature of the subject – whether to his sensibility (inclination and taste) or to his understanding and reason, which by the special constitution of their nature employ themselves with delight^l upon an object – it would, strictly speaking, be nature that gives the law; and this, as a law of nature, must not only be cognized and proved by experience – and is therefore in itself contingent and hence unfit for an apodictic practical rule, such as moral rules must be – but it is *always only heteronomy* of the will; the will would not give itself the law but a foreign impulse would give the law to it by means of the subject's nature, which is attuned to be receptive to it.

An absolutely good will, whose principle must be a categorical imperative, will therefore, indeterminate with respect to all objects, contain merely the *form* of volition as such and indeed as autonomy; that is, the fitness of the maxims of every good will to make themselves into universal law is itself the sole law that the will of every rational being imposes upon

^k *Lehrbegriffe*

^l *Wohlgefallen*

itself, without having to put underneath it some incentive or interest as a basis.

How such a synthetic practical proposition is possible a priori and why it is necessary is a problem whose solution does not lie within the bounds of metaphysics of morals, and we have not here affirmed its truth, much less pretended to have a proof of it in our power. By explicating the generally received concept of morality we showed only that an autonomy of the will unavoidably depends upon it,^m or much rather lies at its basis. Thus whoever holds morality to be something and not a chimerical idea without any truth must also admit the principle of morality brought forward. This section then, like the first, was merely analytic. That morality is no phantom – and this follows if the categorical imperative, and with it the autonomy of the will, is true and absolutely necessary as an a priori principle – requires a possible *synthetic use of pure practical reason*, which use, however, we cannot venture upon without prefacing it by a *critique* of this rational faculty itself, the main features of which we have to present, sufficiently for our purpose, in the last section.

4:445

^m *anhänge*, perhaps “is attached to it”