

## Can We Interpret Kant as a Compatibilist about Determinism and Moral Responsibility?

*(forthcoming in the British Journal for the History of Philosophy)*

Ben Vilhauer

### I. Introduction

According to traditional readings of Kant's theory of free will, Kant is at heart an incompatibilist about determinism and moral responsibility. Incompatibilists hold that there is a basic conflict between determinism and moral responsibility. Compatibilists, on the other hand, hold that there is no basic conflict between determinism and moral responsibility.

On initial inspection, Kant's theory might appear compatibilistic, because he holds that determinism is true, but that we are nonetheless morally responsible. It is usually thought, however, that Kant holds this position because of the distinction in his ontology between agents *qua* phenomena, and *qua* noumena. Kant clearly holds that determinism is true for agents as they appear in time (i.e. agents *qua* phenomena), but not as they are in themselves (i.e. agents *qua* noumena), and that agents *qua* noumena somehow determine the empirical causality of agents *qua* phenomena. The traditional view is that the independence of agents *qua* noumena from the deterministic phenomenal causal series can make room for alternative possibilities of action, and that these alternative possibilities of action are necessary if agents are to be held morally responsible, in a way that compatibilists do not accept. Thus it is thought that Kant uses the metaphysics of the agent *qua* noumenon to undercut the significance of determinism for moral responsibility.

But what if Kant could be interpreted as a compatibilist? That is, what if the traditional interpretation is mistaken, and at the end of the day Kant really thinks there is no basic tension between determinism and moral responsibility? Such an interpretation is advanced by Hud

Hudson, in his book *Kant's Compatibilism*.<sup>1</sup> The purpose of this discussion is to describe some features of Hudson's interpretation, to highlight a reason why this interpretation is important not only for Kant's theory of free will, but also for transcendental idealism as a whole, and to point out some problems for this interpretation.

A compatibilist interpretation of Kant has much to recommend it, in the eyes of those who find much to recommend itself in Kant's epistemology and ethics, but demand a deflationary reading of things in themselves. That is because it is arguable that it is in an incompatibilist reading of Kant's theory of free will that a 'positive' conception of things in themselves does its most important work. Throughout the history of the critical reception of Kant's work, many Kant commentators have felt that the idea of things in themselves which stand 'outside' the empirical world, inaccessible to theoretical reason, but which must nonetheless play a fundamental explanatory role in ontology, epistemology, and ethics, was profoundly confused. Such commentators have thought either that this 'positive' idea of things in themselves could not really have been endorsed by Kant, or that he did endorse it and it was his great error.

These commentators have often sought to interpret Kant in a way that shows the central insights of his epistemological and ethical projects to be independent of this 'positive' conception of noumena. But according to the traditional interpretation of Kant's theory of free will, it relies heavily on such a 'positive' conception, because of the idea that agents *qua* noumena determine the empirical causality of agents *qua* phenomena. Indeed, in Kant's theory of free will, Kant seems to depend more directly on a 'positive' conception of noumena to solve a philosophical problem than at any other point in his philosophical work. But if we can make do with a compatibilist interpretation of Kant's theory of freedom, we can rid ourselves of the

need for a 'positive' conception of noumena here. That is because we only need to appeal to a 'positive' conception of noumena if we try to reconcile incompatibilism with Kant's obvious commitments to moral responsibility as well as determinism. If we instead suppose that there is no conflict between the deterministic necessitation of agents *qua* phenomena and moral responsibility, there is no need to undercut the significance of determinism by appealing to agents *qua* noumena which stand 'outside' the deterministic causal series. And if the traditional reading is in error, and Kant is truly a compatibilist, then there is indeed no conflict to resolve. Since one might reasonably suppose that if we can do without a 'positive' conception of noumena here, we can do without it throughout Kant's philosophy, this is an important issue not just for the interpretation of Kant's theory of free will, but for the interpretation of transcendental idealism as a whole.

## II. Hudson's Interpretation

The guiding idea in Hudson's interpretation is that Kant's theory of freedom can be interpreted in terms of Donald Davidson's anomalous monism<sup>2</sup>. And there are indubitably points of similarity between anomalous monism and transcendental idealism. Anomalous monism analyzes agents in terms of 'two aspects' in a way that may strike one as similar to Kant's 'two aspect' analysis of the agent in terms of the empirical and intelligible character. On anomalous monism, physical events and mental events are the two 'aspects', and their relationship is such that they are in a sense different, but have an underlying unity, and this may also strike one as similar to Kant's empirical character/intelligible character distinction.

For Davidson, mental and physical events are token-identical but type-distinct. Davidson's ontology is thoroughly monistic regarding substance, so that all events are physical

events, i.e. every mental event is also a physical event. But in the explanation of mental events, no physical types occur, and in the explanation of physical events, no mental types occur.

Physical events are necessitated according to deterministic laws, which means that mental events are too, in a sense, since all mental events are physical events too. But events are only covered by laws insofar as they instantiate types contained in the laws, and Davidson holds that causal laws only contain physical types. Thus mental events, *qua* the mental types in terms of which they are individuated, are nominally independent of determinism. Davidson also holds that to be a cause is to be subsumed by a strict, i.e. exceptionless and deterministic, causal law. So mental events are only causes insofar as they are identical with physical events, because strict laws only subsume events under physical types. Davidson thinks this leaves room for ordinary ascriptions of causality to mental states, of the kind ‘my desire for an apple caused me to pick an apple’; Davidson just holds that, if this statement is true, it will be so by virtue of a strict law connecting the intentionality-laden events of desiring the apple and picking the apple *under their physical descriptions*. Davidson holds that we have no reason to suppose that we will ever be able to explain mental events in physical terms, and thus bring mental types into the system of types related according to deterministic laws, because on his view the holism of the mental makes it such that our identification of mental types is always subject to revision. In other words, since the process of interpreting human behavior is fundamentally open-ended, we can never have a system of mental types that is fixed once-and-for-all, which we could attempt to bring into correlation with the deterministic system of physical types.

Hudson thinks that the kind of independence from determinism which mental events have in anomalous monism is sufficient for the kind of independence from determinism required by the intelligible character in Kant’s account of free will. The purpose of interpreting Kant’s

theory of freedom this way is to show the standpoint of practical reason to be fundamentally distinct from that of theoretical reason, as Kant insists it to be, and show that internal to the standpoint of practical reason there is a non-physicalistic mode of intentional explanation that is crucial to our understanding of our place in the world, as Kant thought there was, without having to appeal to a special kind of noumenal causation, or to appeal to a metaphysics of things in themselves that stand behind empirical objects and out of sight of theoretical reason.

Hudson imports Davidson's notion that mental events can be seen as causal in the ordinary way, because they are identical with physical events which are covered by strict laws. In Hudson's interpretation, such mental causation becomes Kant's causality of reason. In this way, Hudson makes room for the idea that practical reason determines the will, without admitting any kind of ontological conditions for empirical objects other than empirical causal conditions.

Hudson also holds that, given Kant's view that causal necessitation of states is a condition for the possibility of determining their place in the order of time, the fact that mental events *qua* mental events are not causally necessitated enables them to model the atemporality of the intelligible character. This allows Hudson to provide an account of atemporality without allowing that there are things which exist atemporally. Mental events, being token-identical with physical events, can be temporally determined through their identity with physical events—it is only when individuated only in terms of mental types that they are not temporally determinable. This interesting result on its own, even apart from the other parallels Hudson draws, would demonstrate the value of studying Kant's theory of free will in light of anomalous monism.

### III. Comparing Hudson's Interpretation to Kant's Texts

Hudson's account is a careful reconstruction of Kant's theory which is sensitive to many aspects of Kant's texts. But it is not clear that the kind of independence mental events have in anomalous monism is sufficient to capture all the ways in which Kant thinks of the free agent as independent from the deterministic causal series. Let us begin evaluating Hudson's interpretation in light of the passage which is the greatest hurdle for a such a reading: the 'freedom of a turnspit' passage, where Kant argues against what he takes to be Leibniz's account of free will, but where his target is broad enough that it would seem to include all compatibilist accounts of free will:

Suppose I say of a human being who commits a theft that this deed is, according to the natural law of causality, a necessary result of determining grounds in the preceding time, and that it was therefore impossible that it could have been left undone. How, then, can appraisal in accordance with the moral law make any change in it and suppose that it could have been left undone because the law says that it ought to have been left undone? That is, how can that man be called free at the same point of time and in regard to the same action in which and in regard to which he is nonetheless subject to an inexorable natural necessity? It is a wretched subterfuge to seek to evade this by saying that the *kind* of determining grounds of his causality in accordance with natural law agrees with a comparative concept of freedom. According to this comparative concept, we can call something a free effect, if its determining natural ground lies within the acting being, e.g., that which a projectile accomplishes when it is in free motion, in which case we can use the word 'freedom' because while it is in flight it is not impelled from without...in the same way the actions of the human being, although they are necessary by their determining grounds which preceded them in time, are still called free because the actions are caused from within, by representations produced by our own powers...[In] the question about that freedom which must be placed at the basis of all moral laws and the imputation appropriate to them, it does not matter whether the causality determined according to a natural law is necessary through determining grounds lying inside the subject or outside him...if, as is admitted by [the advocates of the comparative concept of freedom] these determining representations have the ground of their existence in time and indeed in the antecedent state, and this in turn in a preceding state, and so on, so that these determinations may be internal and may have psychological instead of mechanical causality, that is, produce actions by means of representations and not by bodily motions: they are always determining grounds of the causality of a being insofar as its existence is determinable in time and thus

under the necessitating conditions of past time...and they therefore leave no transcendental freedom, which must be thought as independence from everything empirical, and thus from nature in general, whether it is regarded as an object of inner sense in time only or also of outer sense in both space and time; without this transcendental freedom...no moral law is possible and no imputation in accordance with it. For this reason, all necessity of events in time according to the natural laws can be called the 'mechanism of nature'...Here we refer only to the necessity of natural law, whether the subject in which this development takes place is called *automaton materiale*, when the machinery is impelled by matter, or with Leibniz *spirituale*, when it is impelled by representations; and if the freedom of our will were merely the latter (psychological and comparative but not also transcendental, i.e. absolute), then it would at bottom be nothing better than the freedom of a turnspit, which, once wound up, also accomplishes its motions of itself. (2C5:95-5:97)<sup>3</sup>

Hudson comments on this passage as follows:

In this passage Kant argues against what he takes to be a Leibnizian theory of freedom, but careful attention to his text shows that what he says is not also a rejection of the view I have attributed to him. What he objects to in this passage is the claim that freedom obtains when 'the determining natural cause is internal to the acting thing.' On the present reconstruction, though, freedom is not accounted for in terms of natural causes at all, whether internal or external to the acting thing. Rather it is to be found in the applicability of an *intelligible* cause, independent of natural causes. Again, he objects to the view of freedom's consisting in being determined by reason, 'if these determining conceptions themselves have the ground of their existence in time or in the antecedent state.' But the present reconstruction denies just this. Nondetermining descriptions of the sort appropriate for the pro-attitude, propositional determination of the causality of reason do not yield temporal determination of their objects, and therefore no 'antecedent state' or law is such as to determine the causality of reason. (52)

(A few words of explanation: by 'nondetermining descriptions of the sort appropriate for the pro-attitude, propositional determination of the causality of reason', Hudson means descriptions involving what in Davidson's account are mental types, and in his reading of Kant, descriptions of the practical determinations of the will.) Hudson focuses on the part of the passage where Kant argues against the idea that freedom obtains when 'the determining natural cause is internal to the acting thing', and in response argues that his theory is not vulnerable to this argument. But it is also important to bear in mind Kant's overarching concern in the

passage. To understand this overarching concern, we must look at the beginning, where he says the following: ‘If I say of a human being who commits a theft that this deed is, in accordance with the natural law of causality, a necessary result of determining grounds in preceding time, then it was impossible that it could have been left undone; how then, can appraisal in accordance with the moral law make any change in it and suppose that it could have been omitted because the law says that it ought to have been omitted?’ In other words, if we say an immoral action is causally necessitated, then how can we suppose that the agent could have avoided doing it, as we must when we evaluate the agent morally? What Kant says is a ‘wretched subterfuge’ is the effort to avoid this question by saying that the immoral action is the effect of deterministically necessitating natural causes ‘within’ the agent, i.e. psychological causes, rather than material causes.

In the excerpted passage, Hudson explains that his account, though compatibilistic, is not affected by this point, because the causality of reason is distinct from the causality of natural law, and ‘nondetermining descriptions’, independent as they are from the time-order determined by causal laws, do not cause their effects as antecedent states. There may be reason to doubt this claim, because it would seem that, for Hudson, reason is only causal at all because of its identity with physical events, and physical events *do* cause their effects as antecedent states. But even if Hudson can make out this claim, it does not provide a direct answer to Kant’s question about how the compatibilist can make sense of alternative possibilities.

Hudson does not appeal to the anomalous monism element in his interpretation to explain alternative possibilities, and it may be useful to consider why it cannot do so. Mental events are identical to physical events, which are deterministically necessitated. Mental events, *qua* mental, are not deterministically necessitated, but this does not open up enough room to claim that, in

any sense which would be significant for moral responsibility, the event even *qua* mental could have been left undone. All it leaves room for is the claim that, individuated in mental terms, the event could have had a different significance. According to the doctrine of anomalous monism, because of the holism of the mental, the ways we identify, individuate, and relate mental types are always subject to revision, so with (for example) different cultural norms of interpretation, two physically identical events could legitimately have different mental descriptions. Thus an event *qua* mental could have been different in the sense that it could have had a different significance in a different mental context. But this kind of alternative would not seem to have much in common with what Kant means by ‘left undone’.

Hudson has an independent argument to which he appeals in order to handle alternative possibilities. Hudson does not argue that alternative possibilities are *not* required for moral responsibility, as many contemporary compatibilists do, following Harry Frankfurt.<sup>4</sup> Hudson instead joins a venerable tradition of compatibilists in claiming that we can still make sense of alternative possibilities if determinism is true.<sup>5</sup> According to traditional incompatibilistic interpretations of Kant, Kant thinks we can make sense of alternative possibilities if determinism is true, but *only* because agents *qua* noumena have alternative possibilities between which they freely choose in determining the empirical causality of agents *qua* phenomena, alternative possibilities which are *not* constrained by the deterministic structure of the empirical causal series, since determinism is only true for appearances, not for things in themselves. Hudson rejects this: on his account, agents *qua* noumena *are* constrained by the deterministic structure of the empirical causal series. It is just that, when we describe agents in practical terms, our descriptions do not involve references to their deterministic necessitation. Hudson claims this does not rule out the alternative possibilities necessary for moral responsibility.

In the following excerpts, Hudson examines the claim that an agent could have done otherwise in terms of the claim ‘An agent could have rendered it false that x.’ That is, Hudson frames the claim that a free agent could have acted differently than he in fact acted in terms of a statement ‘x’ about an action than an agent took, but might not have taken. Thus for the agent to have acted differently would have been for the agent to have rendered it false that x.

Suppose, for example, that we are asked to consider the following pair of potential definitions for [the] problematic phrase [‘an agent could have rendered it false that x’], one causal and one noncausal reading:

(D1) An agent could have rendered it false that x=df : There is some action the agent could have performed such that, if the agent had performed that action, then either that action or a causal consequence of that action would have falsified x.

(D2) An agent could have rendered it false that x=df : There is some action the agent could have performed such that, if the agent had performed that action, then something would have falsified x. (93-4)

If determinism is true, affirming ‘An agent could have rendered it false that x’ on D1 would entail that the agent has causal power over the past, or over the causal laws determining x. That is, the agent could have acted in some way that caused the past or the causal laws to be different. Hudson thinks that human beings could not have either such causal power. Certainly it seems safe to say that no coherent metaphysics could hold that we have the power to causally affect the past. And it must also be acknowledged that, from the perspective of a physicalistic metaphysics, such as the Davidsonian one at the foundation of Hudson's interpretation of Kant's transcendental idealism, it would be absurd to suppose that we have the power to causally affect the laws of nature.

It must be noted, however, that according to the traditional interpretation of Kant's theory of free will, it is not absurd to suppose that we have the power to causally affect the laws of nature. From this perspective, we can hold that both the causally necessitated phenomena of the empirical world, and the particular empirical causal laws which necessitate them, are mere appearances of things in themselves, and that some things in themselves are human agents.

Choices of maxims by agents *qua* noumena are the ontological substrates of both (1) the empirical-psychological events which constitute the choices of agents *qua* phenomena and (2) the particular causal laws that necessitate those empirical-psychological events. If choices of maxims by agents *qua* noumena had been different, then they would have had different appearances—that is, the empirical-psychological events which constitute the choices of agents *qua* phenomena would have been different, and the particular causal laws necessitating them would have been different too.<sup>6</sup>

Given Hudson's approach, however, he reasonably enough rules out D1, and offers an account of 'An agent could have rendered it false that x' in terms of D2, according to which we have a kind of non-causal control over the events. Hudson explains that D2

simply says that [the agent] could have done otherwise than perform [the action taken] at t, and if he had done otherwise, then either the past would have been different than it was or the laws would have been different than they are; [it] does not say that [the agent] has any causal power over the past or any causal power over the laws. (94-95)

This idea requires careful interpretation. At first pass, the point may seem to be that what it means for an agent to have alternative possibilities on D2 is just for it to be the case that the agent *does* do otherwise in alternative possible worlds with different pasts or different causal laws, i.e. that this is a sufficient condition of an agent's being able to do otherwise. But this cannot be an answer to Kant's overarching concern in the 'freedom of a turnspit' passage, i.e. the question 'if we say an immoral action is causally necessitated, then how can we suppose that the agent could have avoided doing it, as we must when we evaluate the agent morally?'

Presumably, all theories of agency will be able to make room for the idea that agents do otherwise in alternative possible worlds, including the theories Kant criticizes in the passage. So if Kant had accepted this account of alternative possibilities, he should have thought that the

theories he criticizes as unable to accommodate alternative possibilities *can* in fact accommodate alternative possibilities.

The D2 compatibilist is more cautious than this, however. The D2 compatibilist is not committed to advancing a sufficient condition of the ability to do otherwise, but only to *rejecting* a *necessary* condition of the ability to do otherwise, that is, rejecting that it is a necessary condition of an alternative possibility of action that so acting is consistent with the actual past and actual laws.<sup>7</sup> But the D2 compatibilist is *too* cautious to satisfactorily respond to Kant's overarching concern. Kant wants to know how the agent could have done otherwise if determinism is true. But the D2 compatibilist only states that the agent could have done otherwise, and if he had done otherwise, then either the past would have been different, or the laws would have been different. Surely Kant would want to press the question, 'but *how* could the agent have done otherwise, if so doing would have required the past or laws to be different?' If the traditional interpretation of Kant's theory of free will is right, Kant can answer this question by positing agents *qua* noumena which are independent of the deterministic empirical causal series. He can hold that agents *qua* noumena establish the particular causal laws determining the empirical causality of agents *qua* phenomena when they choose maxims. This is because the particular causal laws necessitating the choices of agents *qua* phenomena, along with the empirical-psychological events constituting those choices, are nothing more than the spatiotemporal appearances of the choices of maxims by agents *qua* noumena.<sup>8</sup> But if D2 compatibilism only rejects a necessary condition, and does not propose a sufficient condition, then it does not provide an answer that could replace the one available within the traditional interpretation. If we *do* try to imagine a sufficient condition for alternative possibilities on D2, we would seem to be driven either toward one which looks a great deal like the one suggested

above, which Kant could not accept, or toward one which involves something like the causal power over the past or laws entailed by D1, which Hudson rejects.

Another difficulty for Hudson's interpretation, closely connected with the one just considered, is as follows. On Hudson's interpretation, the intelligible character can be a cause only via its identity with physical events, or in Kant's terms its identity with the empirical character, where all the 'real' causation happens, so to speak. So for Hudson, we must explain intelligible causation in terms of the causation of the empirical character according to natural laws, but it would be wrong to suppose that, in explaining the causation of the empirical character according to natural laws, we had to refer to the intelligible character. But Kant appears to do just this. The following passages are examples:

If we grant that effects are appearances and that their cause is also appearance, is it necessary that the causality of their cause must be merely empirical? Could it not instead be, that although every effect in appearance requires a connection with its cause according to laws of empirical causality, this empirical causality, without any interruption of its connection with natural causes, is itself an effect of a causality that is not empirical but intelligible? A544/B572

a rational being can rightly say of every unlawful action he has done that he could have left it undone, even though as appearance it is sufficiently determined in the past and, so far, is inevitably necessary; for...he imputes to himself, as cause independent of all sensibility, the causality of those appearances. 2C5:98

insofar as the acting person regards himself at the same time as noumenon...he can contain a determining ground of that causality in accordance with laws of nature, and this determining ground of natural causality is itself free from all laws of nature. 2C5:114

Hudson's interpretation does not offer any way of accommodating this idea. On Hudson's compatibilistic account, there is no reason to *want* to appeal to the agent *qua* noumenon in explaining the causation of the empirical character according to natural laws, because there is no need to undercut the significance of determinism for moral responsibility and make room for incompatibilist-style alternative possibilities. But when we see the pains Kant takes to argue for

the possibility of this kind of explanatory connection, we have another reason to doubt that Hudson's interpretation can accommodate Kant's own conception of alternative possibilities.

It must be emphasized that Hudson clearly and thoroughly documents compatibilistic leanings in Kant's earlier works, and it is possible that, through his completion of *Groundwork of the Metaphysics of Morals*, Kant was in fact a compatibilist (although, even in the *Groundwork*, Kant seems clear that we must believe we have incompatibilist-style free will insofar as we are agents, despite the possibility of interpreting him as saying that we must doubt such freedom insofar as we are theoretical reasoners.) But Kant's mature moral philosophy, as presented in the Second Critique (where the 'freedom of a turnspit' passage appears) and afterwards, is much more difficult to interpret compatibilistically. The account of alternative possibilities which Hudson advances may well work for Kant's moral philosophy prior to the Second Critique, but it does not seem able to do so afterwards.

To sum up, it seems very unlikely that the recalcitrant aspects of Kant's texts considered here could be accommodated within any interpretation that seeks a deflationary account of agents *qua* noumena. Unless we posit noumena which are more independent from the empirical causal series than Hudson allows, the deterministic necessitation of agents *qua* phenomena will constrain agents *qua* noumena too, so agents will only have alternative possibilities in a compatibilistic sense which appears unacceptable to Kant. So it looks like we may be unable to provide a complete account of Kant's theory of free will without a metaphysics of transcendent noumena that is certain to seem baroque to many philosophers. However, Hudson's interpretation remains noteworthy as a careful, clear reconstruction of Kant in light of some of the most

important accounts of agency produced by the analytic tradition, which repays close study and reflection.

*Claremont McKenna College*

## NOTES

---

<sup>1</sup>Hud Hudson, *Kant 's Compatibilism*, Cornell University Press, Ithaca 1994.

<sup>2</sup> Hudson draws elements of his account from Ralf Meerbote, who may have been inspired by Davidson's own work. See, e.g., Ralf Meerbote, 'Kant on the Nondeterminate Character of Human Actions', in *Kant on Causality, Freedom, and Objectivity*, ed. W.L. Harper and Ralf Meerbote, University of Minnesota Press, Minneapolis 1984. See also 'Mental Events', in Davidson, Donald, *Essays on Actions and Events*, Clarendon Press, New York 1980.

<sup>3</sup> References to Kant's texts will be made as follows: material from the first *Critique* will be cited by page in A and B editions. Second *Critique* material will be cited as "2C". Texts used are as follows: *Kritik der reinen Vernunft*, Felix Meiner Verlag, Hamburg 1998; *Kritik der praktischen Vernunft*, Felix Meiner Verlag, Hamburg 1990. Translations are my own, in consultation with the following translations: *Critique of Pure Reason*, Norman Kemp Smith, St. Martin's, New York, 1929, and Werner Pluhar, Hackett, Indianapolis, 1996; *Critique of Practical Reason*, Lewis White Beck, Macmillan, New York, 1985, and Mary Gregor, Cambridge, New York 1997.

<sup>4</sup> See e.g. Frankfurt's "Alternate Possibilities and Moral Responsibility", *Journal of Philosophy*, 1969; 66, 828-39.

<sup>5</sup> The approach to alternative possibilities Hudson offers here can be seen as one of a school which has unquestionably been influential in the history of free will theory, which holds that determinism is not in conflict with alternative possibilities. One branch of the school argues for what is often referred to as a 'conditional analysis' of the ability to do otherwise. It may begin in Hume's account, in *An Enquiry Concerning the Human Understanding*. (*Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, ed. L.A. Selby-Bigge, Oxford, 1975.) Hume defines freedom—liberty, in his terms—this way: 'By liberty...we can only mean a power of acting or not acting according to the determinations of the will; that is, if we choose to remain at rest, we may; if we choose to move, we also may.' Hume's definition may, at first glance, appear to leave room for alternative possibilities of a kind which should be acceptable to Kant. But in context, it is clear that, though Hume is endorsing the idea that we might have chosen otherwise, and that if we had chosen otherwise, we would have acted differently, he is *not* endorsing the idea that it was *in our power* to have chosen otherwise. The background thought is that if something had been different about the past, then we might have been *causally necessitated* to choose otherwise. So we could have done otherwise if the past or causal laws

---

had been different, but we had no power over the past or causal laws. G.E. Moore defended a position similar to Hume's, and he is widely credited with introducing it into contemporary analytic philosophy. Its core is the same as that of Hume's account— Moore holds that one *could* have done otherwise just in case one *would* have done otherwise *if one had chosen* to do otherwise, while tacitly acknowledging that one was causally determined to choose as one did. There are contemporary defenders of the conditional analysis as well, but Hudson is not among them: he belongs to a diverging branch of this school, 'altered-past'/'altered-law' compatibilists, which holds that we can appeal to a notion of 'non-causal power' over the past and/or the causal laws to avoid the counterintuitive consequences of the conditional analysis, but still reconcile determinism with alternative possibilities, without positing agents which stand beyond the deterministic empirical causal series. Defenders of this approach include David Lewis (e.g. "Are We Free to Break the Laws?", *Theoria*, 1981; v.47, 113-121). John Martin Fischer also has a helpful discussion of the view in his "Introduction" to *Moral Responsibility*. (*Moral Responsibility*, ed. John Martin Fischer, Cornell University Press, Ithaca, 1986.)

<sup>6</sup> I examine this interpretation in detail in 'Transcendental Freedom and Causal Laws in Kant', forthcoming.

<sup>7</sup>For this point, thanks are due to Hud Hudson (personal correspondence).

<sup>8</sup>See note 6.