

## The Scope of Responsibility in Kant's Theory of Free Will

*Benjamin Vilhauer, William Paterson University of New Jersey*

### Abstract

In this paper, I discuss a problem for Kant's strategy of appealing to the agent qua noumenon to undermine the significance of determinism in his theory of free will. I then propose a solution. The problem is as follows: given determinism, how can some agent qua noumenon be 'the cause of the causality' of the appearances of that agent qua phenomenon without being the cause of the entire empirical causal series? This problem has been identified in the literature—Ralph Walker provides what is perhaps the most dramatic presentation of it. But it has never received an adequate solution. In this paper, I argue that Walker's objection is only decisive if we must understand our responsibility as responsibility for events, but not causal laws. I argue that we need not interpret Kant's theory in this way. I demonstrate that each agent qua noumenon could be responsible for "limited instantiation scope" causal laws which necessitate only the phenomenal actions of that same agent qua phenomenon. Part of this project involves showing that there are relevant constituents of actions which are "rare" enough to instantiate such laws. I demonstrate that, on Kant's view, events in agents' bodies are not rare enough, but events in agents' phenomenal souls are.

### I. Introduction

Kant's mature moral philosophy is incompatibilistic about determinism and moral responsibility.<sup>1</sup> Incompatibilists hold that there is a basic conflict between determinism and

---

Acknowledgements: I am indebted to many people for conversations and comments which helped me in refining the interpretation presented here, especially Michael Forster, Robert Pippin, Candace Vogler, Ted Cohen, Eric Watkins, Allen Wood, Simon Saunders, Timothy Rosenkoetter, Rachel Zuckert, and the anonymous reviewers at the *British Journal for the History of Philosophy*.

moral responsibility, while compatibilists hold that there is no such conflict. Upon initial inspection, Kant's theory might appear compatibilistic, because he holds both that determinism is true, and that we are morally responsible. Kant thinks that the truth of determinism is demonstrated by the Second Analogy, that is, by the conclusion that the necessitation of all alterations according to causal laws is a condition for the possibility of the experience of objective succession. But he also thinks we are immediately aware that we have incompatibilist-style moral responsibility, an awareness which is based on what he describes in the second *Critique* as a 'fact of pure reason' (*ein Faktum der reinen Vernunft*).<sup>2</sup>

---

<sup>1</sup> It is possible to interpret some passages in the first *Critique* and the *Groundwork* as compatibilistic, but from the second *Critique* forward, compatibilistic interpretations are difficult to sustain.

<sup>2</sup> Kant first presents his critical theory of free will in the Third Antinomy of the first *Critique*, where the idea of a noumenal foundation for empirical causation is presented as a resolution of reason's demand for an absolute completeness or totality of the regressive series of causal conditions. Kant claims that we can imagine how agents might spontaneously originate causal series despite constituting, *qua* phenomena, parts of a regressive series of empirical conditions, if we think of them as noumenal foundations for those parts of the series. But without appealing to the idea, rooted in the Second Analogy, that empirical causal conditions must *necessitate* subsequent determinations, we do not get the threat to moral responsibility which concerns Kant in all his discussions of free will. Further, as Kemp Smith has noted, even if we accept the idea that reason's demand for a completion of the series of causal conditions is resolved by appealing to a noumenal foundation, this idea alone can at best lead us to posit a *single* noumenal foundation for the single global causal series of nature, not individual noumenal foundations for individual phenomenal agents. (Norman Kemp Smith, *A Commentary to Kant's Critique of Pure Reason*, Atlantic Highlands: Humanities Press International, 1992, pp. 517-18.) But this more complex picture is integral to Kant's view of free will. An inference to this picture is only possible if we begin from the 'fact of reason'. Thus for the sake of philosophical clarity, it makes most sense to see Kant's theory of free will as depending directly on the Second Analogy and the 'fact of reason', rather than on the more general conceptions at play in the Third Antinomy. It should be emphasized that a 'strong' interpretation of the Second Analogy is endorsed in this paper, i.e. an interpretation according to which the Second Analogy implies that all events are governed by causal laws. Some commentators, among them Henry Allison, who will be discussed later in this paper, reject this, and instead adopt a 'weak' interpretation according to which the Second Analogy implies particular relations of causal necessitation, but not causal laws. But as an interpretation of Kant, this view is difficult to sustain, in view of Kant's A91/B94 claim that the concept of cause 'strictly requires that something, A, should be such that something else, B, follows from it *necessarily* and in accordance with an *absolutely universal rule*.' For more on this issue, see Michael Friedman, 'Causal Laws and the Foundations of Natural Science', in *The Cambridge Companion to Kant*, edited by Paul Guyer (Cambridge, 1992).

Kant thinks he can hold this position because of his distinction between agents *qua* phenomena, and *qua* noumena: agents *qua* noumena are 'outside' the deterministic empirical causal series, so to speak, and they ground the causality of agents *qua* phenomena. Kant thereby makes room for agents to spontaneously originate empirical causal sequences, and to have alternative possibilities of action. He holds that such freedom is necessary if agents are to be held morally responsible, a view which compatibilists reject. Kant's goal in this is to preserve what he takes to be our everyday incompatibilist conception of moral responsibility, despite the truth of determinism. Said differently, Kant uses the metaphysics of the agent *qua* noumenon to undermine the significance of determinism for moral responsibility. According to Kant, his theory makes it possible for every agent to

say of every unlawful action he perpetrated, that he could have omitted it, although as appearance it is sufficiently determined in the past and, so far, is inevitably necessary; for...he attributes to himself, as cause independent of all sensibility, the causality of those appearances. (2C98)<sup>3</sup>

---

<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', third *Critique* as '3C', *Groundwork of the Metaphysics of Morals* as 'G', *Metaphysical Foundations of Natural Science* by 'MN', all followed by Akademie pagination (i.e. as paginated in *Kants gesammelte Schriften*, hrsg. von der Deutschen Akademie der Wissenschaften, 29 vols., Berlin: Walter de Gruyter, 1902-). Texts used are as follows: *Kritik der reinen Vernunft*, hrsg. von Jens Timmerman (Hamburg: Felix Meiner Verlag, 1998); *Kritik der praktischen Vernunft*, hrsg. von Karl Vörländer, (Hamburg: Felix Meiner Verlag, 1990); *Kritik der Urteilskraft*, hrsg. von Heiner F. Klemme, (Hamburg: Felix Meiner Verlag, 2001). Translations are my own, in consultation with the following translations: *Critique of Pure Reason*, Norman Kemp Smith (New York: St. Martin's, 1929) and Werner Pluhar (Indianapolis: Hackett, 1996); *Critique of Practical Reason*, Lewis White Beck (New York: Macmillan, 1985) and Mary Gregor (Cambridge, 1997); *Critique of Judgment*, Werner Pluhar (Indianapolis: Hackett, 1987) and J.C. Meredith (Oxford Clarendon, 1952) and J.H. Bernard (New York: Hafner, 1951); *Groundwork of the Metaphysics of Morals*, Mary Gregor (Cambridge, 1998) and Lewis White Beck (Indianapolis: Bobbs-Merrill, 1959); *Metaphysical Foundations of Natural Science*, James Ellington (*Philosophy of Material Nature*, Indianapolis: Hackett, 1985).

In the first *Critique*, Kant describes the distinction between agents *qua* noumena and *qua* phenomena in terms of a distinction between 'empirical character' and 'intelligible character':<sup>4</sup>

In a subject belonging to the sensible world we should have first an empirical character, whereby its actions, as appearances, stand in thoroughgoing connection according to constant laws of nature with other appearances. The other appearances constitute conditions from which the actions can be derived, and therefore the actions are, in conjunction with them, members of a single series in the order of nature. Second, one would still have to grant the subject an intelligible character, through which it is indeed the cause of those same actions as appearances, but which does not itself stand under any conditions of sensibility, and is not itself appearance. We can also call the first the character of the thing in appearance, and the second its character as thing in itself. (A539-40/B567-8)

In this paper, I will discuss a problem for this strategy of appealing to the agent *qua* noumenon to undermine the significance of determinism. I will then propose a solution. The

---

<sup>4</sup>I insist upon the clumsy locutions 'agents *qua* noumena' and 'agents *qua* phenomena' in order to highlight the fact that the interpretation to be developed here does *not* assume the existence of two sets of ontologically independent entities, i.e. phenomenal agents on the one hand, and noumenal agents on the other. It posits only one set of ontologically subsistent entities, i.e. agents *qua* noumena. The appearances of those agents *qua* noumena, i.e. agents *qua* phenomena, are ontologically dependent on agents *qua* noumena, and also upon the intuition of human minds. Kant's framing of the distinction as one between two 'characters' has prompted some commentators to suppose that he posits two sets of ontologically independent agents, introducing a host of problems about the relation between these agents, and in particular, about which of the two sets contains the agents who are really morally responsible. (See, e.g., Lewis White Beck, *A Commentary on Kant's Critique of Practical Reason*, Chicago: University of Chicago Press, 1960, and 'Five Concepts of Freedom in Kant,' in *Philosophical Analysis and Reconstruction*, ed. J.T.J. Szrednick, Dordrecht: Martinus Nijhoff, 1987, 35-51.) But the broader thrust of Kant's remarks on this topic suggests that this dual-ontology interpretation is mistaken. Consider 2C95:

if one still desires to save [freedom], no way remains other than to ascribe the existence of a thing insofar as it is determinable in time, and also its causality in accordance with the law of natural necessity, to appearance only, and to ascribe freedom to the *same being* as a thing in itself. [my italics]

It must be emphasized, however, that the interpretation to be developed in this paper is not a version of the interpretation commonly purveyed under the label 'the two-aspect interpretation' (e.g. Henry Allison's interpretation—see Allison 1983, 1990 and 1996). That interpretation rejects the ontological priority of noumena, so it cannot use the ontological priority of noumena to undermine the significance of determinism for moral responsibility, and it therefore cannot accommodate Kant's incompatibilism. The present interpretation can be a 'two-aspect' interpretation just as much as Allison's, so long as it is clear that the noumenal 'aspect' of agents is ontologically prior to the phenomenal 'aspect' of agents. This issue is discussed in more detail in the main text below.

problem is as follows: given determinism, how can some agent *qua* noumenon be 'the cause of the causality' of the appearances of that agent *qua* phenomenon, without being the cause of the entire empirical causal series? This problem has been identified in the literature, but has never received an adequate solution. A solution is significant in its own right, as a matter of Kant interpretation. But it is also significant for contemporary free will theory, because the solution demonstrates in some detail that libertarian-style alternative possibilities can be made compatible with determinism, and this should be of current interest, given the recent upsurge in work on libertarian approaches to free will.

This paper is structured as follows. I will begin by explaining the problem to be addressed in more detail. Next, I will outline the solution to be developed here, and I will describe its connections to previous commentary. Then I will present a more detailed account of the solution.

The problem to be addressed in this paper has been described most dramatically by Ralph Walker, though it is pointed out earlier by Kemp Smith.<sup>5</sup> Walker claims that

my noumenal self is responsible not just for the particular decision I made on Thursday afternoon, but for the entire series of causes which constitute my empirical character and which led up to that decision. And we can hardly stop here. Since it belongs to the phenomenal world my empirical character must also be causally determined— by heredity, upbringing, and the effects other people have had on me; if my responsibility is to be salvaged these also must have been freely decided upon by my noumenal self. Indeed, in view of the thoroughgoing causal interaction that the Third Analogy requires, my noumenal self must have freely chosen the entire causal series that makes up the phenomenal world...my responsibility extends far beyond my own character: I can be blamed for the First World War, and for the Lisbon earthquake that so appalled Voltaire.<sup>6</sup>

Walker's objection is flawed in one respect, but he nonetheless puts his finger on a deep problem for Kant. The problem for Kant is as follows. Kant's goal is to incorporate

---

<sup>5</sup> Kemp-Smith, 1992, 517-18.

<sup>6</sup> Ralph C.S. Walker, *Kant* (Boston: Routledge & Kegan Paul, Boston, 1978) pp. 148-9.

alternative possibilities of action into the deterministic empirical causal series, so that agents can be the first causes of their actions despite determinism. But it would appear that, if determinism is true, then our actions could have been different only if the entire causal series had been different. In other words, given determinism, it seems that we can only be the first causes of our own actions if we are the first causes of all the events in the deterministic causal sequence. If this problem cannot be solved, then the consequences of Kant's metaphysics of free will must be profoundly counterintuitive. If Kant's theory can only secure our responsibility by grossly distorting its scope, then it is not a viable theory.

The flaw in Walker's objection is in his claim that, on Kant's theory, we must be *morally* responsible for events far beyond our own phenomenal characters, as illustrated by his claim that one might be blameworthy for the First World War. Walker misses a distinction between moral responsibility and causal responsibility. Any theory which makes agents the first causes of their actions, and which attempts to accommodate our customary views about morality, must recognize that there are some events for which we have causal responsibility but no moral responsibility, that is, events which we neither intend to cause nor are negligent in causing. For example, if a passerby rescues a child from a burning building, and the child grows up to commit a crime, the passerby has partial causal responsibility for the crime but no moral responsibility. This is because she did not intend to causally contribute to the crime, and had no way of knowing that she would be doing so. So, suppose Walker is right that on Kant's theory we can only be the first causes of our own actions if we are also the first causes of all the causal antecedents of our actions, including events in the distant past before our own births. Presumably we neither intend to cause these events nor are negligent in causing them. So we might well understand our responsibility for those events as equivalent to the passerby's responsibility for the crime, that is, causal

responsibility without moral responsibility. It might be argued that this is sufficient to defend Kant against Walker.

It is not sufficient, however, for the core of Walker's objection still stands, i.e. the objection that Kant's theory distorts the scope of our responsibility in a way that prevents it from accommodating our everyday conception of moral responsibility. Determinism appears to imply that one can only be the first cause of one's own actions by being the first cause of all one's actions' causal antecedents. This is because the causal antecedents of one's actions, along with the laws of nature, make one's actions causally necessary, and it is hard to see how one could be the first cause of one's actions unless one was also the first cause of anything that made them causally necessary. But it will commonly be the case that one agent's actions have the actions of other agents as causal antecedents. (In fact, given the high degree of social interaction characteristic of human life, this will probably be the case almost universally.) The real threat of Walker's objection is that, in such cases, it appears that one can only be the first cause of one's own actions by being the first cause of those other agents' actions. Since Kant holds that one cannot be morally responsible for one's actions unless one is the first cause of those actions, this would seem to imply that, in cases where one's actions have the actions of other agents as causal antecedents, one can only be morally responsible for one's actions if those other agents are *not* morally responsible for their actions. What Walker's objection shows, therefore, is that Kant's theory threatens to distort our everyday conception of the scope of our moral responsibility not by expanding it, as Walker himself claims, but rather by diminishing it, since the moral responsibility of some agents would come at the expense of the moral responsibility of other agents. (Since the core of Walker's objection stands unanswered, however, I will continue to refer to this problem as 'Walker's problem'.) To solve Walker's problem, we must demonstrate that Kant's theory of free will can in fact accommodate our everyday conception of the scope of our causal responsibility,

and with it, our everyday conception of the scope of our moral responsibility. In what follows, I will sometimes describe Walker's problem simply as a problem about the scope of our responsibility, but the distinction between the direct problem for causal responsibility, and the indirect problem for moral responsibility, must be borne in mind.

## II. Connections to Previous Commentary

To begin developing a solution to Walker's problem, let us consider Allen Wood's account of Kant's theory of free will, and his remarks defending Kant against Walker, in his much-noted paper 'Kant's Compatibilism'. According to Wood,

A particular timeless choice of my intelligible character affects the natural world by selecting a certain subset of possible worlds, namely those including a certain moral history for my empirical character, and determining that the actual world will be drawn from that subset of possibilities. (90-91)<sup>7</sup>

On this interpretation, we have reason to suppose that any particular agent *qua* noumenon determines something less than the entirety of the empirical causal series. But, as Wood goes on to acknowledge in subsequent remarks, this restriction is not strong enough to set the bounds of our responsibility where we customarily take them to be set, since it is not clear that one could select such a subset of possible worlds without determining events far outside what we customarily take to be the scope of our responsibility. Wood argues, however, that this is not a problem for Kant's theory:

Even if my choice somehow issues in a world containing the First World War, the Lisbon Earthquake, and the deeds of Idi Amin, it seems reasonable to hold me morally responsible only for those events which must belong to the actual course of things because I have the empirical character or fundamental maxims that I do. Kant must admit that on his theory this may include events that happen at places and even times remote from my life history in the temporal world. Yet Kant can reply that because in principle we know nothing about how our timeless choices operate on the temporal world, it must be impossible for us to say with confidence which events these may be. It

---

<sup>7</sup>Allen W. Wood, 'Kant's Compatibilism', in *Self and Nature in Kant's Philosophy*, edited by Allen W. Wood (Ithaca: Cornell University Press, 1984).



seems open to Kant to suppose that they correspond to those events for which we normally regard ourselves as morally responsible.(92)

In this passage, Wood proposes what are really two different responses to Walker's problem. The first is that, even if my choices do wind up determining events far outside what we customarily take to be the scope of my responsibility, it is only reasonable to hold me morally responsible for the events which must be part of the causal series for me to have the empirical character or fundamental maxims that I have. In other words, if I could have the empirical character and fundamental maxims I *actually* have in some *other* possible world in which my choices did *not* issue in the deeds of Idi Amin, then it is not reasonable to hold me morally responsible for the deeds of Idi Amin. This claim is not obviously false, and it might constitute a satisfactory defense if the real threat posed by Walker's problem was that the scope of our moral responsibility might be wider than we customarily take it to be. But, as argued in the previous section, the real threat is quite different—it is rather that there might not be enough moral responsibility to go around. The sheer fact that my choices made the deeds of Idi Amin inevitable would not imply that I was morally responsible for his deeds, but it would quite clearly imply that he was *not* morally responsible for his deeds. What is needed is an interpretation of Kant where moral responsibility is not (as it were) a scarce resource for which agents have to compete.

The second response Wood proposes is quite different: since 'we know nothing about how our timeless choices operate on the temporal world', it is 'impossible to say with confidence which events' we are causally responsible for, so it 'seems open to Kant to suppose that they correspond to those events for which we normally regard ourselves as morally responsible'. A natural way of interpreting this idea is as follows. Because of the limitations on our theoretical knowledge of ourselves *qua* noumenal agents, there is no way for us to know whether or not we are causally responsible for events outside what we customarily take to be the scope of our causal responsibility. This makes it consistent to

suppose that the events for which we are causally responsible are just the events for which we customarily take ourselves to be causally responsible, thereby making room for the claim that we are all the first causes of our own actions, and for the commitment to free will which is entailed by the 'fact of reason'. But when confronted with the claim that determinism can be squared with our customary views about the scope of our causal responsibility, reason demands more than mere consistency. Reason demands an *explanation* of how it can be true, and this is why Kant presents a metaphysics of free will in addition to his account of noumenal ignorance. The problem for Wood is that his interpretation of Kant's metaphysics does not provide an adequate explanation, because he gives no argument to show that it is possible to determine our empirical characters by selecting subsets of possible worlds without determining events outside what we customarily take to be the scope of our responsibility.<sup>8</sup> We need an interpretation that sets narrower bounds on our responsibility if we are to solve Walker's problem.

The first step in providing such an interpretation is to distinguish between two elements of a deterministic causal series: first, the determinations at the various points in time in the causal series, and second, the causal laws which necessitate the relations between determinations and antecedent and subsequent determinations. Given determinism, the determinations at any one time, along with the laws, necessitate all other determinations at antecedent and subsequent times. Therefore, control over the determinations at any one time

---

<sup>8</sup>Wood might be read as holding that that no argument *can* be given to show that it is possible to determine our empirical characters by selecting subsets of possible worlds without determining events outside what we customarily take to be the scope of our responsibility. The interpretation of Kant's theory of free will presented here will (indirectly) provide such an argument. It will do this by way of demonstrating the possibility of causal laws which are instantiated only within particular empirical characters. This means that Wood's interpretation would be viable if supplemented by the analysis presented in this paper. But this interpretation of Kant's theory does not draw on the idea of selecting subsets of possible worlds, and this gives it an advantage over Wood's, since this idea does not appear in Kant's own descriptions of noumenal agency.

would entail causal responsibility for the determinations at all points in time in the causal series. So establishing causal responsibility for the agent *qua* phenomenon in this way would entail responsibility for the rest of history. Walker appears to assume that the agent *qua* noumenon must exercise control over the agent *qua* phenomenon in something like this fashion.

Let us therefore suppose that the agent *qua* noumenon determines the agent *qua* phenomenon by controlling causal laws, rather than by directly controlling determinations. This sort of approach has been discussed recently in independent publications by the present author and Eric Watkins.<sup>9</sup> According to Watkins, 'Kant's solution involves the selection of laws of nature, which depend on the natures of things',<sup>10</sup> and 'natures are, in turn, constituted by the essential grounds of substances'.<sup>11</sup> Watkins holds that the key insight in Kant's theory of free will is that 'by exercising its causal powers, a substance might be able to choose (some aspect of) its own nature, which influences in turn which laws of nature hold and thus which events are necessary in accordance with them'.<sup>12</sup> More specifically, Watkins interprets the

---

<sup>9</sup> Eric Watkins, *Kant and the Metaphysics of Causality* (New York: Cambridge University Press, 2005). Ben Vilhauer, 'Can We Interpret Kant as a Compatibilist about Determinism and Moral Responsibility?' *British Journal for the History of Philosophy*, 12(4)2004: 719-730, and also *An Interpretation and Defense of Kant's Theory of Free Will* (Ph.D. Thesis, University of Chicago, 2002). This approach may be anticipated by a remark Jonathan Bennett makes in a response to Wood's paper (Jonathan Bennett, 'Kant's Theory of Freedom', in *Self and Nature in Kant's Philosophy*, ed. Allen W. Wood (Cornell, Ithaca, NY, 1984). Bennett criticizes Wood's apparent acceptance of the idea that one's 'timeless choice must affect the whole course of the world's history'.<sup>9</sup> He suggests that

Perhaps Kant himself would try in some less fatal way to reconcile my empirical character's being naturally caused with my freely choosing to have it. That may be what he is doing when he describes as 'an effect of intelligible causality' not *this causal chain* but *this empirical causality*' (B572; italics added), suggesting that what I freely choose are not the causally interrelated items but rather the causal relation that links them. (104)

<sup>10</sup> Watkins 2005, p. 303. Related points are discussed in Vilhauer 2002 and 2004.

<sup>11</sup> Watkins 2005, p. 346.

<sup>12</sup> Watkins 2005, p. 361.

agent *qua* noumenon as a noumenal substance which partially determines its own nature, thereby setting some of the laws of nature and making it the case that that same agent *qua* phenomenon is determined in particular ways.

Watkins' rationale for supposing that we are not 'completely responsible for the laws of nature and everything that follows from them (including all the evils that transpired throughout the course of history)' <sup>13</sup> is as follows:

It is important to understand that the laws of nature are nothing other than laws of the natures of things. That is, the laws of nature that hold in a given world are a function of the natures that are instantiated in that world. Thus, it is important to stress that what personal agents freely choose are not immediately the laws of nature but rather their own natures. Once all the natures in the world have been determined (whether through free choice or otherwise), then the laws of nature (and the necessity that depends upon them) are set.<sup>14</sup>

As Watkins puts it elsewhere, on his interpretation, agents 'immediately choose their own natures and only indirectly the laws of nature that are based on them'.<sup>15</sup> Said differently, the content of an agent's choice is that agent's nature, not the laws of nature. Nonetheless, we can explain the effect of that agent's choice on the phenomenal world in terms of determining laws of nature, because the laws of nature are just a function of the natures instantiated in that world. In other words, the laws of nature are just a function of the natures of the substances which belong to that world.<sup>16</sup>

---

<sup>13</sup> Watkins 2005, p. 335. Watkins does not cite Walker as the source of this problem, but given that Watkins addresses this problem during a discussion of Wood's response to Walker, it is clear that Walker is in fact the source.

<sup>14</sup> Watkins 2005, pp. 335-6.

<sup>15</sup> Watkins 2005, p. 334. This point is also discussed in Vilhauer 2004, p. 727. Also see note 25 below.

<sup>16</sup> Watkins holds that noumenal substances belong to a world insofar as they bear causal relations to one another. As Watkins puts it, 'Kant takes [it] to be an analytical claim...that noumenal substances can belong to a single world in virtue of their (causal) relations to each other' (Watkins 2005, 353).

Watkins' interpretation takes an important first step toward a solution of Walker's problem, but it leaves us some distance to go, because it is not clear that it is possible to determine the laws which govern one's own actions without also determining the laws which govern the actions of some other agents.<sup>17</sup> This is because it is not clear how general, and how frequently instantiated, we must suppose the laws that govern our actions to be. Without a more detailed account of how noumenal choices determine laws of nature, we cannot secure the moral responsibility of one agent without jeopardizing the moral responsibility of other agents.<sup>18</sup>

Before developing this more detailed account, however, let us forestall some very general objections to the claim that we have some choice about what causal laws obtain. At

---

<sup>17</sup> The interpretation advocated here (and in Vilhauer 2002 and 2004) overlaps with Watkins' interpretation in some fundamental respects, but it carries the analysis of Kant's theory of free will farther.

<sup>18</sup> Derk Pereboom makes a related point: he states that Watkins 'fails to specify the mechanism whereby the phenomenal laws are fixed by the noumenal choices' (though he does not discuss this issue in the context of Walker's problem). According to Pereboom, the mechanism can be specified by adopting Luis de Molina's view that

God knows, eternally, what every possible libertarian free creature would choose in every possible circumstance, and with this knowledge, God is able to direct the course of history with precision, partly in virtue of creating just those free creatures whose choices fit a preconceived divine plan. On a version of this Molinist view adapted to Kant's idealism, God would reconcile noumenal transcendental freedom with phenomenal determinism by creating just those transcendently free beings the appearances of whose free choices conform to the deterministic laws that God intends for the phenomenal world.

(Pereboom, 'Kant on Transcendental Freedom', forthcoming in *Philosophy and Phenomenological Research*.)

Pereboom's interpretation solves Walker's problem by bypassing it entirely. But there is a problem with supposing this to be Kant's own view. Pereboom's interpretation cannot incorporate Kant's idea that agents *qua* noumena are the 'causes of the causality' of their phenomenal actions unless it is supposed that those agents are overdetermining causes. That is, if God sets the laws for the phenomenal world, and creates just those particular transcendently free agents whose choices conform to the causal structure of the phenomenal world He has designed, then God would seem to be a sufficient cause of those agents' phenomenal actions. In light of this, we can only suppose that the agents are also the causes of their phenomenal actions if we suppose that they are overdetermining causes.

first glance, such a claim may sound absurd. Choosing the laws of nature is, after all, a power traditionally reserved for God.

If we accept transcendental idealism, however, it is entirely natural to suppose that human agents are responsible for causal laws. Kant holds that we cannot have theoretical knowledge of noumena, but this is often taken to imply a much broader exclusion of knowledge about noumena than is warranted by his texts. Theoretical knowledge is knowledge of determinations, i.e. synthetic knowledge that particular predicates apply to things, and this is only possible regarding objects in space and time. But we can know that noumena exist without synthetically knowing that any particular predicates apply to them, if the existence of noumena is implied by basic, general features of transcendental idealism of which we have a priori knowledge.

Kant thinks that from the synthetic apriority of our knowledge of space and time, we can conclude that space and time are transcendently ideal, i.e. that they are contributed to empirical reality by the human mind. But we also know that what we thereby contribute is only a formal feature of reality, a blank manifold of empty spatiotemporal extension. The empirical objects which make up the specific content of empirical reality, and are known only a posteriori, cannot be wholly constituted by the human mind. Their empirical content must be contributed by something independent of the human mind. If there were nothing independent of the human mind to stand as the ground of the specific content of empirical reality, empirical reality would never amount to more than an empty manifold. Since we know that spatiotemporality is dependent on the human mind, we must isolate, in philosophical reflection, non-spatiotemporal ontological substrates which are the grounds of the specific content which appears in spatiotemporally extended empirical objects. These ontological substrates are noumena. To borrow an expression from Paton, noumena provide empirical content through their contribution of the 'particularity' of the properties of empirical

objects.<sup>19</sup> That is, noumena explain the fact that empirical objects instantiate the particular, contingent properties they instantiate rather than some others.

Those who reject the view that noumena are ontological substrates for phenomena may object that empirical causal laws can explain why empirical objects instantiate the properties they instantiate.<sup>20</sup> But such an explanation cannot be complete, because it does not tell us why the particular causal laws which obtain are the laws they are, rather than different laws. Noumena are needed to explain the particularity of causal laws in the same way that they are needed to explain the particularity of the properties of empirical objects.

Understanding and intuition together are responsible for constructing the objective temporal order of the world, and (as Kant explains in the Second Analogy) they do this by imposing the form of deterministic causal necessitation on all empirical events. But what is thereby imposed only amounts to the formal, general fact *that there are deterministic causal laws*. This a priori construction is not responsible for the fact that the particular causal laws which obtain are the laws they are, rather than different laws. That is the role of noumena. This point allows us to make sense of the idea of human responsibility for causal laws. Since we ourselves are noumena as well as phenomena, we are responsible for some of the laws of nature. Kant's theory of free will is based on the supposition that the laws for which we are responsible are the laws that necessitate our actions *qua* phenomena.

It must be emphasized, however, that the sense of 'responsibility' according to which we are responsible for causal laws is a derivative one. Our free choices *qua* noumena are

---

<sup>23</sup> Paton notes 'the distinction, upon which Kant always insists, between empirical and universal laws and between empirical and universal concepts. Only what is strictly universal is imposed by the mind upon objects. Empirical differences are particular determinations of the universal, but their particularity is not due to the mind and must be due to things.' (H.J. Paton, H.J. *Kant's Metaphysic of Experience*, London: George Allen & Unwin, 1936, Vol. 1, p. 139.)

<sup>20</sup> Thanks to Robert Pippin for this objection.

choices of maxims, i.e. choices of the principles we act upon, not choices of causal laws. Our choices of maxims appear in inner sense as temporally extended phenomena of empirical psychology, necessitated by causal laws. The practical types and laws under which we choose our maxims and the theoretical a posteriori types and laws under which these choices appear to us are entirely different, and as a consequence of our theoretical ignorance of noumena, we cannot know *why* they correlate as they do. But as a consequence of the fact of reason, we are committed to the belief that they *do* in fact correlate, in such a way that if our choices of maxims had been different, the empirical-psychological events which are their appearances would have been necessitated according to different causal laws.<sup>21</sup>

It is clear that any interpretation developed along these lines cannot be a version of what is called the 'two-aspect' interpretation in contemporary Kant commentary. To suppose that noumena are the grounds of the particularity of the properties and causal laws of phenomena is just to suppose that noumena are ontologically prior to phenomena. But denying the ontological priority of noumena is the very *raison d'être* of what is called the 'two-aspect' interpretation. Throughout the history of Kant commentary, some commentators have seen the idea of noumena which transcend space and time as part of the peculiar majesty of the Critical Philosophy, but others have seen it as a metaphysical monstrosity. Commentators of the latter persuasion have thought either that this idea was not actually endorsed by Kant, or that it was his great error; they have often proposed deflationary accounts of transcendental idealism which endeavor to avoid positing the existence of such transcendent noumena while preserving the central insights of the Critical Philosophy. What

---

<sup>21</sup> This explanation is based on one given in Vilhauer 2004 (p. 727). Also see Watkins 2005, p. 335, for a discussion of similar points.



is called 'the two-aspect interpretation' is the latest flowering of this persuasion, and Henry Allison is perhaps its most influential proponent in contemporary Kant scholarship.<sup>22</sup>

Allison rejects 'the 'noumenalistic' view that grants ontological priority to things as they are in themselves',<sup>23</sup> and he holds that transcendental idealism is not committed to the existence of any non-spatiotemporal things.<sup>24</sup> That is, according to his account of the non-spatiotemporality of noumena, the only things under consideration in transcendental idealism are empirical objects, which exist in the manifold of space and time constructed in intuition. When we think of things as they are in themselves, independent of their relation to human sensibility, we think of empirical objects in abstraction from their mind-dependent spatiotemporality, but this way of representing empirical objects does not correspond to any underlying ontological reality. As Allison puts it, 'Kant's transcendental distinction is primarily between two ways in which things (empirical objects) can be 'considered' at the metalevel of philosophical reflection'—to 'consider things as they are in themselves is to

---

<sup>22</sup> See Allison 1983, 1990 and 1996. Also see note 4 above.

<sup>23</sup> Allison 1996, p. 11.

<sup>24</sup> Interpretations like the one to be developed here, which do accept the existence of non-spatiotemporal things, needn't ignore the 'neglected alternative' problem, i.e. the problem that it does not follow from the fact that space and time as we experience them are the product of the human mind that there is not also a mind-independent spatiotemporal manifold (or at least a mind-independent manifold of extension which is analogous in important ways to space and time) which is occupied by things in themselves. To make room for agents qua noumena with alternative-possibility freedom, it is enough to know that noumena do not occupy the particular spatiotemporal manifold constructed through the determination of our form of intuition according to schematized categories. The crucial point for Kant's theory of free will is not that agents qua noumena are atemporal, but rather that they do not occupy the temporal manifold constructed a priori by the human mind, because it is only if they occupy the temporal manifold we construct that we must represent them as deterministically necessitated. We know that deterministic necessitation is a condition for the possibility of objective succession in the temporal manifold we construct, but we have no reason to believe that this is true for other possible kinds of temporal manifolds which are not based on the form of intuition of finite, passively intuiting beings like ourselves. Indeed, we can hold that the 'fact of pure reason' entails that deterministic necessitation is *not* a condition for the possibility of objective succession in any temporal manifold occupied by things in themselves.

reflect on them in a way which ignores or abstracts from the subjective conditions of human sensibility'.<sup>25</sup> That is, the only things to be considered in transcendental idealism are empirical objects, and while we can consider them in abstraction from their spatiotemporality, we are thereby merely considering unequivocally spatiotemporal objects in abstraction from their spatiotemporality.<sup>26</sup>

If we follow Allison in interpreting transcendental idealism as not committed to the existence of non-temporal things, we find ourselves in a terrible bind in interpreting Kant's theory of free will. The Second Analogy implies that all objects in time are deterministically necessitated. But if we follow Allison in rejecting the ontological priority of things in themselves, we cannot suppose that atemporal agents qua noumena serve as the ontological substrates of agents qua phenomena, and shape the empirical causal structure of agents qua phenomena to make room for free will. We are left without any way to undermine phenomenal determinism. This means there are fundamental obstacles to accommodating Kant's incompatibilism within Allison's interpretation.<sup>27</sup>

---

<sup>25</sup> Allison 1996, p. 3.

<sup>26</sup> Rae Langton glosses what she calls 'Allison's deflationary proposal' this way: 'When doing philosophy, we sometimes consider things in abstraction from their relation to our sensibility, in abstraction from their spatial [and] temporal...properties; but this does not show that there are non-spatial, atemporal...things' (Langton 1998, p. 9).

<sup>27</sup> Allison is clearly committed to an incompatibilist interpretation of Kant, claiming, for example, that

[A]t the heart of Kant's account of freedom in all three *Critiques* and in his major writings on moral philosophy is the problematic conception of transcendental freedom, which is an explicitly indeterminist or incompatibilist conception (requiring an independence of determination by all antecedent causes in the phenomenal world). (Allison 1990, p. 1)

The problem for Allison is that, given the basic structure of his interpretation of transcendental idealism, the only way for him to accommodate the incompatibilism he endorses is to weaken Kant's phenomenal determinism. The most dramatic manifestation of this is his rejection of Kant's claim that human actions are, in principle, predictable. One example of Kant's predictability claim is at A550/B578:

[All] the actions of a human being in appearance are determined...according to the order of nature, and if we could investigate all the appearances of men's

But rejecting Allison's version of the 'two-aspect' interpretation does not entail accepting what is commonly called a 'two worlds' interpretation in contemporary Kant commentary, i.e. an interpretation according to which noumena and phenomena constitute two sets of ontologically independent entities. The interpretation to be developed here can claim to be a 'two-aspect' interpretation just as much as Allison's, since it posits only one set of ontologically subsistent entities, i.e. noumena. The difference is that the interpretation to be developed here insists upon the ontological priority of noumena. Phenomena, i.e. empirical objects as well as particular causal laws, are merely relational properties of noumena, though in a special and complex sense. Phenomena are constituted by second-order relations between noumena and human intuition, i.e. by *relations of the relations* between noumena and human intuition.<sup>28</sup>

First-order relations between noumena and human intuition are the relations whereby our intuition is passively affected by noumena. These relations make up the purely sensible content of intuition, transcendently prior to the determination of experience according to the schematized categories. This purely sensible content 'fills in' various spatiotemporal locations in the blank manifold of pure intuition, but it is indeterminate: it is what Kant calls 'intuitions without concepts', and describes as 'blind'. It is not experience, and no empirical objects or laws are to be found within it. For human minds, experience requires the

---

wills to their grounds, there would not be a single human action we could not predict with certainty and recognize as necessary from its antecedent conditions.<sup>27</sup>

Allison rejects this, stating that 'Kant has neither the need nor the right to assert...that, given sufficient knowledge, we could infallibly predict human actions' (Allison 1983, p. 326). If one's goal is to provide an interpretation of Kant, rather than a reconstruction, then Allison's interpretation on this point must be recognized as robbing Peter to pay Paul. (Also see note 39 below.)

<sup>28</sup> Kant himself of course does not distinguish between first- and second-order relations between noumena and human intuition, but this distinction is a helpful way of foregrounding an important point which is implicit in Kant's texts.

cooperation of the forms of intuition and schematized categories. Empirical objects and laws are second-order relations *between* these first-order relations. More specifically, empirical objects and laws are relations between the locations in space and time that are 'filled in' with purely sensible content by the first-order relations.<sup>29</sup> These second-order relations are spontaneously constructed through the successive synthesis of the manifold, in which the schematized categories are applied to purely sensible content. In other words, the construction of these second-order relations, the construction of empirical objects and laws, and the application of the schematized categories to purely sensible content, all amount to one and the same spontaneous activity of transcendently constituting empirical reality.

This interpretation explains the ontological foundation of the particular causal laws governing agents *qua* phenomena as follows. The particular causal laws governing agents *qua* phenomena are second-order relations between agents *qua* noumena and intuition. The particularity of these laws, i.e. the fact that the laws which obtain are the laws they are rather than some other laws, is the result of first-order relations between agents *qua* noumena and human intuition that 'fill' locations in both the outer and inner intuitions of agents *qua* phenomena with purely sensible content. The laws have the form of deterministic necessitation because it is imposed upon them as they are constructed according to the schematized category of causality.

---

<sup>29</sup> See, e.g., B67-8: 'Whatever in our cognition belongs to intuition contains nothing but mere relations: of places in intuition (extension), of change of places (motion), and of the laws according to which this change is determined (motive forces). But what is present in that place, or what effect...it produces in the things themselves, is not given to us by [the cognition which belongs to intuition].'

### III. Limited-Instantiation-Scope Laws

Once we have a general account of how transcendental idealism can explain human responsibility for causal laws, we have taken the first step toward solving Walker's problem. But it is only the first step, because, as we saw in considering Watkins' approach, it is not clear that an agent could be responsible for some of the laws of nature, i.e. the laws necessitating her own actions, without being responsible for all the laws of nature. Unless we can demonstrate this to be possible, we have no reason to suppose that one agent's moral responsibility does not come at the expense of other agents' moral responsibility.

To limit the scope of agents' responsibility in the required way, it would be sufficient to demonstrate that an agent *qua* noumenon could be responsible for causal laws which necessitate only her own phenomenal actions, and no other events. This would require causal laws whose instantiation was limited to the scope of an individual phenomenal agent's empirical character. In other words, we would need causal laws which were instantiated only within the spatiotemporal boundaries of an individual phenomenal agent. Let us refer to these laws as limited-instantiation-scope (LIS) laws.

If we explain our responsibility in terms of control over LIS laws, we could represent ourselves as the first causes of our actions without the counterintuitive consequences that Kant's theory of freedom seemed to threaten. We would be left with partial causal responsibility for events in the distant future, but as discussed above, everyday morality also accommodates such partial causal responsibility for events in the distant future. (Remember the passerby who saves the child who grows up to commit a crime.) So this is not an argument against an interpretation based on LIS laws. Further, if we assume that all agents control LIS laws necessitating their phenomenal actions, then we can reject the possibility that some agents are the first causes of other agents' actions. So an interpretation of Kant's

theory of freedom in terms of LIS laws would appear to preserve our everyday conception of the scope of our responsibility.

It might be objected that LIS laws are impossible, because of basic facts about the structure of the empirical causal series. Two features often thought to be fundamental to the empirical causal series are (1) *universal repeated instantiation* of causal laws, and (2) *complete unity* of causal laws. If the empirical causal series had either of these features, it would thwart the idea of LIS laws.

Universal repeated instantiation is the idea that all particular causal laws are repeatedly instantiated at countless points throughout the grand causal series that constitutes nature, thereby establishing the regularity of nature's patterns. If it were necessary for causal laws to be universally repeatedly instantiated in this way, LIS laws would be impossible, because scope-limitation would be impossible.

Complete unity of causal laws is the idea that particular causal laws are unified in such a way that particular laws can always be derived from more general laws. In other words, the idea is that, at bottom, there is only one perfectly general causal law, of which all particular causal laws are the consequences. Complete unity is incompatible with LIS laws because if laws are completely unified, we could only be responsible for any particular law if we were responsible for the general law.

However, Kant's account of causation (as described in the Second Analogy in the first *Critique*) does not entail that causal laws are repeatedly instantiated or unified. All that the Second Analogy entails, and all that we can know a priori about the empirical causal series, is that all events are necessitated according to causal laws. We cannot know anything about the relationships among the particular laws governing events. We cannot know whether, or how often, they are repeatedly instantiated. We also cannot know whether, or to what degree, they are unified. The Second Analogy describes the a priori contributions of the understanding

and the forms of intuition to the empirical causal series. Their contributions are exhausted by the fact *that there are causal laws* necessitating all events. Facts about *which* causal laws there are, and what they are like, e.g. whether they are repeatedly instantiated and how unified they are, are due to the contribution of the particularity of causal laws by noumena. With the significant exception of some a priori knowledge about the laws of physics (to be considered in more detail later) all features of particular causal laws except *that they are causal laws* are grounded in noumena.

We need not claim there is *no* repeated instantiation or unity. On the contrary, there clearly is some unity and repeated instantiation, as manifested in the successes of natural science. We can take unity and repeated instantiation to be two aspects of what Kant calls the systematic unity of empirical laws (in section 4 of the First Introduction to the third *Critique*). But systematic unity is a regulative principle rather than a constitutive one, and thus is not a criterion which causal laws must satisfy in order to count as causal laws. There must be some systematic unity if the understanding is to be able to 'find itself' in nature (3C193), i.e. if it is to be possible for the understanding to orient itself in nature in the ways needed for the development of natural science. Said differently, some unity and some repeated instantiation together constitute a condition for the possibility of natural science. But this is not a problem for the LIS laws approach, for LIS laws in the empirical character would only require a minute divergence from universal repeated instantiation and unity of causal laws, and would leave more than enough systematic unity for natural science to thrive. As we will see in more detail later, this turns out to be significant for Kant's philosophy of science, because he holds that the laws of empirical psychology cannot be systematically unified.

LIS laws would not require us to reject the view that like causes produce like effects, or the view that causal laws are universal in form. For it is possible to have a causal law which is perfectly universal in form, such that all events of type A causally necessitate events

of type B, which is only instantiated once in the course of nature. That is, it is possible that there is only one event of type A in the course of nature.

It might be objected that it is empty to say that a causal law *necessitates* an event in the case of a once-instantiated law. Surely causal necessity involves universal concomitance, at the very least, such that events of type A *always* cause events of type B. But what is the force of the *always* if the types involved occur only once? This objection would be problematic if Kant accepted Hume's inductivism about causal laws, that is, if Kant thought that the very concept of causal law was something we abstracted from our observations of patterns of repeated event-types. If Kant accepted this view, it would be quite problematic for him to suppose that any particular causal law existed without observing repeated instances of it. But of course Kant rejects Hume's view. On Kant's view, our knowledge that all events are causally necessitated is a precondition for our developing an objective conception of the empirical world, since, according to the Second Analogy, causal necessitation is a condition for the possibility of objective succession. If we had to abstract the concept of causal law from our observations of the empirical world, we would never get so far as having an objective conception of the empirical world, *or* having the concept of causal law. Kantian causal laws are counterfactual-sustaining generalizations that are known to obtain independent of any a posteriori knowledge about how many times they are actually instantiated. For us to form an objective conception of an event, we must suppose that the event is the effect of some distinct event which is its cause, and that the cause and effect are subsumable under some pair of types, say A and B, which are such that, in all cases, *if* A is instantiated, *then* B is instantiated. We can explain what it means for A to causally



necessitate B in this way even if there is only one actual A. It simply means that, if there had been another A, it would have caused another B.<sup>30</sup>

#### IV. Locating Instantiations of LIS Laws

At this point, we must focus more finely on the details of empirical reality. Can we really suppose that nature contains only one instance of some cause-types without violating other assumptions about nature that neither Kant nor we would wish to relinquish? If we are to explain phenomenal actions as governed by LIS laws, we must identify relevant constituents of those actions which are *rare* enough to instantiate those LIS laws.

---

<sup>30</sup> Another objection to the LIS laws interpretation of Kant's theory of free will might be advanced on the basis of the *Critique of Teleological Judgement*. In this text, Kant argues that we must invoke final causes to explain the special sort of organization we find in organic unities such as living bodies (including human bodies). One might suppose that in such cases we get a complete causal explanation through final causation, and we need not suppose that the stretches of empirical reality containing organic unities are necessitated according to causal laws at all. (Kant often refers to causation according to causal laws, i.e. the kind of causation that competes with final causation in the third *Critique*, as 'mechanical' causation.) This might lead one to suppose that we can dispense with LIS laws altogether, since final causes can play the role needed in the present theory. But there is clear evidence against such an approach. At the most general level, we should observe that it is not obvious that final causation is a kind of causal necessitation, and the Second Analogy clearly requires causal necessitation if we are to represent succession as objective. Since Kant advances no argument in the *Critique of Teleological Judgement* which would exempt organic unities from this requirement of the Second Analogy, the default interpretation would seem to be that organic unities fall under two kinds of causal principles, final causes as well as (mechanical) causal laws. And this is indeed Kant's position in the *Critique of Teleological Judgement*. Kant holds that causal laws are constitutive causal principles, while final causes are merely regulative. As Kant puts it at 3C235'-6':

In our empirical investigation of nature's causal connection, we can and should strive to [proceed] in terms of nature's merely mechanical laws as far as we can, since in these laws lie the true physical bases for [an] explanation [of nature, the bases] which [in their] coherence constitute...scientific knowledge[.]

Kant's goal in the Solution to the Antinomy of Teleological Judgement is to make final causation compatible with causal laws. Kant makes it clear here that final causes in no way supplant causal laws:

where we think purposes as bases that make certain things possible, we also must assume means whose causal law does not *itself* require anything that presupposes a purpose, in order for this law to be mechanical and yet also a subordinate cause of intentional effects. (3C414)

To begin, let us reflect on free will theory for a moment and ask what sort of constituents of actions could serve as instantiation-points (so to speak) for LIS laws. Kant thinks that agents have alternative possibilities of action despite determinism. We can think of such alternative possibilities as modal branch-points in the causal series, such that if the agent acts in one way, the causal series branches in one direction as it 'travels' into the future, and if the agent acts in a different way, the causal series branches in a different direction. The point of introducing LIS laws is to make it possible to describe events at those alternative-possibility branch-points as necessitated by laws that agents are responsible for *qua* noumena, laws that could have been different, and which therefore could have necessitated that the causal series 'traveled' along a different modal branch if agents had chosen differently. But to make this work, we have to identify plausible events to serve as the branch-points in our explanation.

There are two key elements in the initiation of a free action that we must consider, in sketching the branch-point events. The first element is the empirical process of forming a phenomenal intention to act one way rather than another. Described differently, it is the empirical choice between alternative possibilities of phenomenal intention. This empirical choice occurs in the empirical character, and is merely a psychological event necessitated by empirical-psychological laws—it is to be definitively distinguished from the noumenal choice, which is part of the intelligible character, and which is the ontological substrate of the particular causal laws which necessitate the empirical choice. The noumenal choice is the formation of a noumenal intention, i.e. the determination or application of a maxim. The free noumenal choice appears in inner sense as the necessitated empirical choice.

The second element to consider is the physical initiation of the bodily movement involved in the action. We might think of this as the 'choice' between alternative possibilities of motion. Presumably this is an event in the brain. ('Choice' is in scare quotes because these

alternative possibilities of motion are the phenomenal expressions of noumenal freedom just as much as the alternative possibilities of phenomenal intention are, but the phenomenal consciousness of making a choice is present only in the choice between phenomenal intentions, not in the physical initiation of one of a set of alternative possibilities of motion.)

These two elements give us two sets of events as candidates for instantiating LIS laws. It might initially seem convenient to adopt some species of physicalism about the mind, so we could explain both sets of events as really consisting in one set of events. But Kant is a dualist, at least at the empirical level:

dualism has a place in psychology...in the empirical sense of the term...  
[Matter], as substance in [the realm of] appearance, is actually given to outer sense, just as the thinking I, likewise as substance in [the realm of] appearance, is given to inner sense (A379-80)

Kant's dualism turns out to be crucial for the LIS laws approach, since as we will shortly see, there are insurmountable obstacles to locating instantiations of LIS laws in material substance. According to Kant's philosophy of mind, a human agent must be understood as a complex of two phenomenal substances, one material, and one immaterial. In some places Kant says the substance of the phenomenal mind, as given to inner sense, is not really substance at all, but is rather a kind of quasi-substance thought of as analogous to substance in space. However, Kant never drops his commitment to empirical dualism, and he thinks of this quasi-substance as substantial enough (so to speak) to be a locus of force, as we will later see. So substance in outer sense and 'substance' in inner sense share enough attributes to support a strong notion of dualism, no matter how we interpret Kant on these points. (I will refer to this quasi-substance in inner sense in quotes, as 'substance', to mark these issues.)

The obstacle to applying LIS laws to the material substance of the agent's body comes from a problem about microphysical reduction, which arises in the following way. At first glance, it seems easy to find events in agents' bodies which are rare enough to instantiate LIS laws. Brain events would be certainly be rare enough to instantiate LIS laws, if individuated

in terms of physical segments of brain structure of significant spatial extension. Brains rank among the most complex structures in the world, and with complexity comes diversity: the probability of two brains, or two events within brains, being identical in structure is vanishingly small. But we should be able to explain macrophysical brain states and events in terms of their microphysical constituents, and these microphysical constituents are not rare enough to instantiate LIS laws. From the contemporary perspective, if we descend to the microphysical structural level, the events at the atomic level which make up the cell-level events of neuron firings are not rare enough to instantiate LIS laws, since the same kinds of atomic-level events make up all cell-level events in all bodies throughout the human species, and throughout much of the rest of the biological world as well. Any theory that implied that we were responsible for the causal laws governing all these phenomena would define the scope of our responsibility much too broadly.<sup>31</sup> Though Kant's anti-atomist theory of matter as force is very different from the conception of matter in contemporary science, where the notion of fundamental particles plays a significant role, Kant still has a strong notion of microphysical reductive explanation. Thus, though material bodies do not have *absolute* microphysical parts, they still have *relative* microphysical parts to which we must appeal in explaining their macrophysical properties. In other words, though all parts of material bodies can be indefinitely subdivided further into parts, there are still relative parts which can be stably individuated for purposes of explanation. Kant makes it clear that the motions of macrophysical material bodies can be explained in terms of the motions of their parts in the *Metaphysical Foundations of Natural Science*. This means that the motions of the brain can be decomposed into the motions of its parts for the purposes of explanation. And though

---

<sup>31</sup>This problem is the phenomenal correlate (so to speak) of the problem about the generality of the nature of substance discussed above in connection with Watkins' interpretation. That is, the problem discussed in connection with Watkins' interpretation was about the generality

Kant's chemistry is quite different from contemporary chemistry, he would not have accepted that the kinds of microphysical parts in brains are rare enough to instantiate LIS laws—he would have assumed that the types of microphysical parts in brains are the same as those in the bodies of the rest of the animal kingdom.<sup>32</sup>

In light of this microphysical reduction problem, we cannot suppose that LIS laws apply to material bodies. But perhaps we can suppose that 'imitation' LIS laws apply to material bodies. That is, we might suppose that *unlimited* instantiation scope laws, i.e. laws that are instantiated both within and beyond the empirical character, have relational 'clauses' which simulate LIS laws. If every macrophysical brain state which initiates a free action is unique, as seems likely, then if we assume that we can give microphysical reductive explanations of these states, it follows that each state is composed of a unique array of microphysical parts. That is, the same types of microphysical parts will probably figure in many or all of these macrophysical states, but the precise nature of the relationships among the tokens of the types will be different in each case. So each of these macrophysical brain states would be composed of what we might call a 'unique relational array' of microphysical parts. We can suppose that these unique relational arrays function as special causal contexts, causing the microphysical parts to behave differently than they behave outside those special causal contexts. The special causal contexts formed by the unique relational arrays could then be factored into the causal laws governing the microphysical parts. That is, the causal law governing the behavior of a specific type of microphysical part would contain a clause

---

of the nature of noumenal substance, and the present problem is about the generality of the nature of material phenomenal substance.

<sup>32</sup>This argument could of course be blocked if one thought Kant might hold that there are fundamentally different kinds of microphysical parts in the brains of different people, i.e. that each brain contained some unique kind of matter. This would certainly be metaphysically possible. But it posits a bizarre sort of complexity at the microphysical level in material nature, and a simpler theory (such as the one I will go on to provide) is certainly preferable. So it would weaken Kant's position to ascribe such a view to him.

specifying a special kind of behavior for it in each of the unique relational arrays in which it figured. If, like Hume, we held a strict inductivist account of causal laws, such relational clauses would amount to a mere cheat whereby exceptions to causal laws were claimed not to be exceptions, by making a show of integrating them into the laws. But Kant is not a strict inductivist, so this is no objection to the relational clause approach.

There is, however, another objection to the relational clause approach which seems decisive. The laws containing relational clauses would have to be laws of physics or chemistry, by Kant's lights, and it seems impossible to accommodate them here. Suppose the relational clauses were in the laws of physics, and thus governed the motions of microphysical parts. That is, suppose that macrophysical brain events are macrophysical motions, to be explained in terms of the motions of their microphysical parts. When these microphysical parts were not elements of the relational arrays constituting relevant macrophysical brain events, i.e. when their motions were not being governed by the relational clauses tacked onto their laws, they would have a characteristic pattern of motion, as necessitated by the non-relational base statement of their laws. This characteristic pattern of motion would be explained in terms of the fundamental forces of physics. But when these microphysical parts entered the relational arrays constituting relevant macrophysical brain-events, they would move *differently*, that is, they move in ways *not* explained by the basic forces of physics. We would have to posit new physical forces *ad hoc*, which began to impinge on these microphysical parts when they entered the unique relational arrays. (If we did not postulate additional forces, we would have to suppose that matter suddenly of itself began to resist the fundamental forces of physics, and that would be absurd.) But Kant would hold, as should we all, that it runs against the basic explanatory orientation of physics to posit new forces willy-nilly. Kant holds that it is essential to the methodology of physics to reduce fundamental forces to as small a number as possible. It is not quite as clear how things would

work if we tried to integrate relational clauses into Kantian chemistry, but there is every reason to think that a parallel set of problems would emerge.<sup>33</sup>

### V. LIS Laws and Empirical Psychology

In light of the barriers to relational clauses in the material sciences, let us try applying LIS laws to the immaterial 'substance' of the phenomenal soul. Kant holds that the phenomenal mind, or phenomenal soul, as he prefers to call it, is extended only in time, while the human body is both temporally and spatially extended. Events in the body are necessitated according to the laws of matter, as are all events in space, but events in the phenomenal soul are independent of the laws of matter, since they are not in space. They are instead governed by the laws of empirical psychology (as Kant explains in the *Metaphysical Foundations of Natural Science*). So perhaps we can find a place for LIS laws among the laws of empirical psychology. Phenomenal mental events would seem to be as diverse as brain-events, so types of phenomenal mental events would seem to be as rare as types of brain events, and thus would also be candidates for instantiating LIS laws. Further, the microphysical reduction problem does not arise regarding the phenomenal mind, since immaterial 'substance' cannot be analyzed into substantial parts.

Immaterial 'substance' cannot be analyzed into substantial parts because there is no phenomenal substrate for the merely temporal appearances of the mind in inner sense. That is, there is nothing in inner sense to play a role parallel to the role played by matter for the appearances of physical objects in space. As Kant puts it at A381-2, 'the appearance to outer sense has something constant and enduring that gives us a substratum lying at the basis of the

---

<sup>33</sup> Another objection to the relational clause approach is that it demands extraordinary complexity in the laws governing microphysical parts of macrophysical brain events which initiate free actions. Since some types of microphysical parts will be involved in many such macrophysical brain events, the laws governing those types of microphysical parts would be

mutable determinations', i.e. matter, while 'time, on the other hand, which is the only form of our inner intuition, has nothing enduring, and therefore allows us to cognize only the variation in determinations, but not the determinable object. For in what we call soul, everything is in continual flux and nothing endures'. When we study a material object, we find that even if its properties change radically, even if it changes enough for us to say that it has ceased to exist, we can still identify an underlying material substrate of substantial parts whose properties remain relatively constant through the change, despite changes in their relations. For example, if a granite sculpture crumbles into fragments, the fragments still have all the properties that make them granite, even though they are no longer related as they must be to constitute the sculpture. We can explain why they previously made up a sculpture, and why they no longer do, in terms of the properties of the bits of granite, and the way their relations have changed. Kant thinks we can always reductively explain material objects in this way at indefinitely smaller and smaller scales. In the case of merely temporal objects like the phenomenal soul, however, there is nothing analogous to matter which remains constant through change, so there is nowhere to look for the substantial parts which make reductive explanation possible.<sup>34</sup>

---

encumbered with one relational clause for each such macrophysical brain event to which they contribute.

<sup>34</sup> The differences just discussed between phenomenal material substance and the 'substance' of the phenomenal soul suggests the following (admittedly speculative) resolution of the problem about the generality of the nature of substance discussed earlier in connection with Walker's interpretation. The regulative principle of systematic unity in the third *Critique* gives us a picture of perfected natural science as positing only one basic kind of material substance, and only one causal law to govern it. On this basis, we might assume that the nature of phenomenal material substance is general. The present interpretation of Kant's theory of free will is compatible with this, since it holds that LIS laws are laws of phenomenal mental 'substance', not of matter. The nature of phenomenal mental 'substance' cannot be general in the same sense, since the phenomenal souls of different agents must be governed by irreducibly different laws. If we suppose that there is a correlation between the natures of phenomenal substances and the natures of the noumenal substances which are their ontological substrates, then we could suppose that the natures of the noumenal substrates of



If we incorporate LIS laws into empirical psychology, free actions will still be inexplicable in terms of the fundamental forces of the material sciences. But we can explain free actions without postulating a miraculous resistance to the forces of the material sciences, and also without postulating new forces within the material sciences and thereby violating the methodology of the material sciences, if we instead suppose that not all the forces which exist are forces of matter—that is, if we suppose that the phenomenal soul has non-physical forces of its own by which it causes alterations in the physical world, i.e. in the brain, where the physical initiations of the bodily movements involved in actions occur. And this is just what Kant appears to do:

one of the forces we attribute to the soul is a *vis locomotiva*, because bodily movements do actually arise whose cause lies in the soul's representations of them, but we do this without trying to ascribe to the soul the only manner in which we know motive forces (namely, through attraction, pressure, impact, and hence motion, which always presuppose an extended being). (3C457)

Placing LIS laws among the laws of empirical psychology also allows us to defend the idea of LIS laws against another important objection. If an LIS law is instantiated only once, it will be impossible to learn its nature inductively through experimentation, because that would require the law to be repeatedly instantiated during each iteration of the experiment. But if our only motivation for supposing there were a barrier to induction in empirical psychology was a desire to accommodate LIS laws, this would be suspect, since it would amount to an a priori restriction on science imposed by the demands of Kant's metaphysics of free will. However, Kant gives us independent reasons for believing that there is such a barrier to induction in empirical psychology, based entirely upon grounds of experimental methodology which are irreproachable if we accept Kant's dualism about mind and body.

---

phenomenal material substances are all the same, while the natures of the noumenal substrates of phenomenal mental substances are not.

Kant holds that empirical psychology is constrained by having merely temporal 'substances' as its subject matter, in such a way as to make it impossible to study empirical psychology experimentally. Kant explains this at MN 471:

the empirical doctrine of the soul must always remain still further removed than chemistry from the rank of what may be properly called natural science, since mathematics is inapplicable to the phenomena of inner sense and their laws [*Gesetze*], unless one might want to take into consideration the law of continuity in the flow of this sense's inner changes, but the extension of cognition so obtained would bear much the same relation to the doctrine of body, as the doctrine of the properties of the straight line bears to the whole of geometry. [This inapplicability is due to the fact that] the pure inner intuition in which the soul's phenomena are to be constructed is time, which has only one dimension. The empirical doctrine of the soul cannot approach chemistry even as a systematic art of analysis or as an experimental doctrine, because in it the manifold of internal observation is separated only by mere thought, but cannot be kept separate and be connected again as one likes; even less does another thinking subject submit to our investigations in a way that is appropriate to our intentions, and even the observation itself alters and distorts the state of the object observed. It can, therefore, never become anything more than a historical...systematic natural doctrine of the internal sense, i.e., a natural description of the soul, but not a science of the soul, nor even a psychological experimental doctrine.

This passage appears in the context of a discussion about how, if a 'body of doctrine' is to count as a science, or a 'pure doctrine of nature', some features of the particular causal laws it contains have to be knowable a priori. All of the objects of physics appear in space as well as time, and explanation in physics depends in large part on relations of objects in space and time. And since we have a priori knowledge of the forms of space and time, Kant thinks that some aspects of physics' particular causal laws can be known a priori. The objects of inner sense only appear in time, and Kant thinks that, given temporality alone, we can apply mathematics only to the 'law of continuity' of the changes of inner sense, but not to the particular causal laws which necessitate its alterations. So there is no a priori mathematical way of reasoning about the particular causal laws of empirical psychology.

Kant's remarks here set out his reasons for holding that we cannot experiment on inner sense. We cannot analyze inner sense because we cannot divide it into parts, since it

has no underlying substrate. Also, observing inner sense necessarily alters it, and we cannot separate the alterations produced by observation from what we are trying to observe. If we cannot experiment on inner sense, we cannot learn the particular laws of empirical psychology by induction. As Kant puts it, our observations of inner sense can never yield ‘a psychological experimental doctrine’.

Since a limit to induction is inherent in empirical psychology, the objection that the LIS laws approach must arbitrarily stipulate a limit to induction in empirical psychology is mistaken. LIS laws can be comfortably accommodated within empirical psychology because systematically unified laws could never arise within empirical psychology to conflict with LIS laws. It is also important to note Kant's implicit view in this passage that we can know there are laws of empirical psychology even though we cannot know *which* particular laws these are. This provides strong support for the earlier interpretation of the Second Analogy which is the basis for the idea of LIS laws.<sup>35</sup>

---

<sup>35</sup> The question of the existence of psychological laws marks a central difference between the present account of Kant's theory of free will and Allison's account. As mentioned above in note 31, Allison rejects Kant's claim that human actions are, in principle, predictable, and a key part of Allison's case for this rejection is that there are no laws of empirical psychology. He says it is a 'major exegetical puzzle' how Kant's claim of such predictability can coexist with his 'denial of the possibility of psychological laws':

If reason and its causality do, indeed, exhibit an empirical character, then the study of that character must pertain to the province of empirical psychology... But Kant denies that empirical psychology is a science, insisting that the most it can provide is a 'natural description...but not a science of the soul' (MAN 4: 471;8)...Given [Kant's]...denial of nomological status to the empirical generalizations of psychology, it becomes...difficult to understand how Kant could insist upon a causal determinism for human actions at the phenomenal level. (Henry Allison, *Kant's Theory of Freedom*, New York: Cambridge University Press, 1990, pp. 31-33.)

Allison is right that Kant thinks empirical psychology is not a science, but he is wrong to claim that Kant 'denies nomological status' to empirical psychology. In the text Allison cites here, *Metaphysical Foundations of Natural Science* 471 (cited just above in the body of the present paper) Kant directly refers to the 'laws' (*Gesetze*) of 'inner sense', as the content of empirical psychology. So Allison seems to misread the text he cites as evidence for his rejection of psychological laws, which he uses in turn as a reason to reject Kant's claim that human actions are, in principle, predictable. Kant's

The goal of this paper has been to show that the LIS laws approach provides a satisfactory solution to Walker's problem, i.e. the problem of explaining the scope of our determination of the phenomenal causal series in a way that preserves our everyday conception of the scope of our moral responsibility. To begin, it was argued that Walker's objection is only decisive if we must understand our responsibility as responsibility for events, but not causal laws. It was then argued that each agent *qua* noumenon could be responsible for particular causal laws which necessitate only the phenomenal actions of that same agent *qua* phenomenon. Next, it was asked whether we could identify relevant constituents of actions which would be rare enough to instantiate such limited-instantiation-scope (LIS) laws. We demonstrated that events in agents' bodies were not rare enough, but that events in agents' phenomenal souls were. So we have found a place in the empirical causal series where LIS laws can be plausibly located. This demonstrates that LIS laws can in fact be used to solve Walker's problem, and more generally, that Kant's theory of free will can combine determinism and alternative possibilities without distorting our everyday conception of the scope of our moral responsibility.

*William Paterson University*

#### References

---

point in *Metaphysical Foundations of Natural Science* 471 is not about whether or not there are psychological laws, though he clearly implies that there are. His point is rather that we cannot know enough about them for empirical psychology to count as a science.

- Allison, H.E. *Kant's Transcendental Idealism* (New Haven: Yale University Press, 1983).
- \_\_\_\_\_. *Kant's Theory of Freedom* (New York: Cambridge University Press, 1990).
- \_\_\_\_\_. *Idealism and Freedom* (New York: Cambridge University Press, 1996).
- Beck, L.W. *A Commentary on Kant's Critique of Practical Reason* (Chicago: University of Chicago Press, 1960).
- \_\_\_\_\_. 'Five Concepts of Freedom in Kant,' in *Philosophical Analysis and Reconstruction*, ed. J.T.J. Szrednick (Dordrecht: Martinus Nijhoff, 1987, 35-51).
- Bennett, J. 'Kant's Theory of Freedom', in *Self and Nature in Kant's Philosophy*, edited by Allen W. Wood (Ithaca: Cornell University Press, 1984).
- Friedman, M. 'Causal Laws and the Foundations of Natural Science', in *The Cambridge Companion to Kant*, edited by Paul Guyer (New York: Cambridge University Press, 1992).
- Hudson, H. *Kant's Compatibilism* (Ithaca: Cornell University Press, 1994).
- Kemp Smith, N. *A Commentary to Kant's Critique of Pure Reason* (Atlantic Highlands: Humanities Press International, 1992).
- Langton, R. *Kantian Humility: Our Ignorance of Things in Themselves*. (New York: Oxford University Press, 1998).
- Paton, H.J. *Kant's Metaphysic of Experience* (London: George Allen & Unwin, 1936).
- Pereboom, D. 'Kant on Transcendental Freedom', forthcoming in *Philosophy and Phenomenological Research*.
- Vilhauer, B. 'Can We Interpret Kant as a Compatibilist about Determinism and Moral Responsibility?' *British Journal for the History of Philosophy*, 12(4)2004: 719-730.
- \_\_\_\_\_. *An Interpretation and Defense of Kant's Theory of Free Will* (Ph.D. Thesis, University of Chicago, 2002).
- Watkins, E. *Kant and the Metaphysics of Causality* (New York: Cambridge University Press, 2005).
- Wood, A.W. 'Kant's Compatibilism', in *Self and Nature in Kant's Philosophy*, edited by Allen W. Wood (Ithaca: Cornell University Press, 1984).
- Walker, R.C.S. *Kant* (Boston: Routledge & Kegan Paul, Boston, 1978).