

Kant, Causation, and Freedom

Critical Notice

E. WATKINS. *Kant and the Metaphysics of Causality*. Cambridge: Cambridge University Press 2005.

Now although there is an incalculable gulf fixed between the domain of the concept of nature, as the sensible, and the domain of the concept of freedom, as the supersensible ... yet the latter **should** have an influence on the former, namely the concept of freedom should make the end that is imposed by its laws real in the sensible world; and nature must consequently also be able to be conceived in such a way that the lawfulness of its form is at least in agreement with the possibility of the ends that are to be realized in it in accordance with the laws of freedom. (CPJ 5, 175-6)¹

-
- 1 For convenience, I refer to Kant's works infratextually in parentheses. The citations include both an abbreviation of the English title and the corresponding volume and page numbers in the standard 'Akademie' edition of Kant's works: *Kants gesammelte Schriften*, Königlich Preussischen, ed. (now Deutschen) Akademie der Wissenschaften (Berlin: G. Reimer [now de Gruyter], 1902-). I generally follow the standard English translations, but have occasionally modified them where appropriate. For references to the first *Critique*, I follow the common practice of giving page numbers from the A (1781) and B (1787) German editions only. For references to Kant's *Reflexionen*, i.e., entries in *Kants handschriftlicher Nachlaß* — which I abbreviate as 'R' — I give the entry number in addition to the Akademie volume and page numbers. The translations from the *Reflexionen* are my own. Here is a list of the relevant abbreviations and English translations:

- CPJ: *Critique of the Power of Judgment*. Trans. P. Guyer and E. Matthews. Cambridge: Cambridge University Press 2000.
- CPR: *Critique of Pure Reason*. Trans. P. Guyer and A. Wood. Cambridge: Cambridge University Press 1997.
- CPrR: *Critique of Practical Reason*. Trans. M. Gregor. In *Immanuel Kant: Practical Philosophy*. Cambridge: Cambridge University Press 1996. 133-272.
- GMM: *Groundwork of the Metaphysics of Morals*. Trans. M. Gregor. In *Immanuel Kant: Practical Philosophy*. 37-108.

I Introduction

Quine notoriously distinguished between those interested in the history of philosophy, and those interested in philosophy. From a contemporary standpoint this is somewhat ironic, because it is clear that Quine himself should have paid more attention to the history of the analytic-synthetic distinction.² This points up the important metaphilosophical fact that one of the ways in which the history of philosophy can actually drive the rest of philosophy is by historically re-examining and re-working seemingly settled topics, and showing us that philosophically things really *were not* the way we have uncritically assumed them to be. That in turn can lead us to question unargued assumptions, and even change our view of the conceptual landscape and the relevant possibilities in logical space. Eric Watkins's *Kant and the Metaphysics of Causality* (henceforth KMC) is a particularly beautiful example of this regressive strategy for moving forward philosophically.

The trick, of course, is to pick your targets carefully: they should be central to the mainstream of contemporary philosophy, not marginal. Watkins has certainly done that. The target he has chosen is *the problem of causation*. His three-part aim is, first, to embed Kant's theory of causation in its 18th century pre-Critical and especially *Leibnizian* setting; second, to argue that Kant's Critical theory of causation is *not* in fact a reply to Hume, and that Kant's metaphysics of causation depends as much on the Third Analogy of Experience and the Third Antinomy of Pure Reason as it does on the Second Analogy; and third, that reconsidering Kant's Critical metaphysics of causation from a pre-Critical and rationalist point of view can contribute to a serious reconsidera-

-
- ID: 'On the Form and Principles of the Sensible and the Intelligible World,' generally known as the 'Inaugural Dissertation.' In *Immanuel Kant: Theoretical Philosophy, 1755-1770*. Trans. D. Walford and R. Meerbote. Cambridge: Cambridge University Press 1992. 373-416.
- MFNS *Metaphysical Foundations of Natural Science*. Trans. J. Ellington. Indianapolis, IN: Bobbs-Merrill 1970.
- MM: *Metaphysics of Morals*. Trans. M. Gregor. In *Immanuel Kant: Practical Philosophy*. 353-604.
- OP: *Immanuel Kant: Opus postumum*. Trans. E. Förster and M. Rosen. Cambridge: Cambridge University Press 1993.
- P: *Prolegomena to Any Future Metaphysics*. Trans. J. Ellington. Indianapolis: Hackett 1977.
- PC: *Immanuel Kant: Philosophical Correspondence, 1759-99*. Trans. A. Zweig. Chicago: University of Chicago Press 1967.

2 See R. Hanna, *Kant and the Foundations of Analytic Philosophy* (Oxford: Clarendon/OUP 2001), ch. 3.

tion of the theory of *agent causation* in the contemporary debate about free will.

II Kant and the Problem of Causation

The concept of causation has two faces. The first face is turned towards the dynamics of physical nature, and the second face is turned towards the dynamics of intentional action. Nevertheless since all the intentional agents we are directly acquainted with are also *animals* and thus living organisms, it seems reasonable to hold that the dynamics of physical nature and the dynamics of intentional action are ultimately the same. This remains true despite a common methodological tendency after 1950 to segregate their metaphysics by shunting physical causation and intentional action onto the specialist spurs of the philosophy of physics and action theory. But philosophical problems, like runaway trains, have a way of getting off the tracks neatly laid for them and smashing into one another — and in this case, they collide precisely at the point where the fundamental biology of human intentional agents meets the fundamental physics of inert matter.

So what is the problem of causation? Since at least the 17th century, physical nature has been thought to operate according to universal strict mechanical laws, whether deterministic or statistical. At the same time intentional agency has been held to operate according to psychological and moral principles that require freedom of the will. In turn, it seems that given an actual psychological or behavioral event *E*, freedom of the will with respect to *E* entails either the ability of the subject to have willed or done otherwise than *E* without qualification (the unconditional alternative possibilities principle) or the ability to have willed or done otherwise than *E* *if* the subject had wanted to will or do otherwise than *E* (the conditional alternative possibilities principle). In either case, freedom of the will does not operate merely according to universal strict deterministic or statistical laws. For on the one hand the unconditional alternative possibilities principle directly entails the denial of natural mechanism (which is the thesis that all events in nature are either determined or else indeterministic and strictly governed by statistical laws³), but also seems to misidentify freedom with brute indeterminism — the existence of causeless, lawless, random events. And on the other hand the conditional alternative

3 See H. Bok, 'Freedom and Practical Reason,' in G. Watson, ed., *Free Will* 2nd ed. (Oxford: OUP 2003) 130-66, at 130-1.

possibilities principle, although it is metaphysically consistent with determinism (and thus with natural mechanism), inserts irreducible first-person conscious desire and volition into processes that are supposed to be fundamentally physical — thereby generating the amazingly hard problem of mental causation.⁴

The problem of causation, then, has two parts. First, how can the dynamics of physical nature and the dynamics of intentional agency each be adequately characterized and related to one another? And second, how can natural mechanism and free will be reconciled? This synoptic way of formulating the problem of causation can be traced directly back to Kant's *Critique of Pure Reason*, and in particular to the Second Analogy of Experience and the Third Antinomy of Pure Reason. But as we all know, Kant's views on the problem of causation are nothing but an apriorist spin on Hume's views. So Hume is the giant, and Kant is standing on his shoulders.

Or do we only *think* we know that? The first unargued contemporary assumption that Watkins is questioning in KMC is that Hume's basic analyses of the concepts of causation (as nothing but contingent extrinsic relations between contiguously successive events), of causal laws (as nothing but contingent general regularities covering relations of contiguous succession between tokens of similar types of events), and of our knowledge of them (as nothing but non-rational imaginative projections from repeated experiences of qualitatively similar constantly conjoined tokens of the same event-types) are accepted by Kant at least as starting points. And the second assumption that Watkins is questioning is that in the Second Analogy Kant offers a direct reply to Hume by simply translating Hume's reductive and skeptical empiricist analysis of causation and causal laws into his own apriorist metaphysics of transcendental idealism.

Watkins wants, however, not only to question but also to *reject* both of these assumptions:

Hume's and Kant's ontologies are radically different, and the lack of a shared vocabulary makes it impossible for one either to find a refutation of Hume's position in Kant's explicit terms or to translate Kant's arguments into Humean terms such that the presuppositions required for a refutation to be possible would be satisfied. If Kant is not attempting to refute Hume, then how should his project be understood? From Kant's perspective, once Hume understands events the way he does and undertakes the project of attempting to construct the world out of them, he makes no mistake in inferring that we can be aware of no necessary connections in

4 See, e.g., A. Mele and J. Heil, eds., *Mental Causation* (Oxford: Oxford University Press 1993).

nature and there is no way to refute this view by pointing out some obvious move available within the empiricist framework that Hume simply overlooked. Rather, Kant recognizes that he must pursue a different project on the basis of a completely different set of ontological and epistemological presuppositions, and his account of causality should thus be understood accordingly as the result of his attempt to articulate a very different alternative to Hume's empiricist account. (KMC, 17)

In this way, Watkins wants to demonstrate that Kant's metaphysics of causality is primarily influenced by rationalist doctrines in the tradition of Leibniz, and only secondarily influenced by empiricist doctrines in the tradition of Hume. Then, having shown how Kant's theory of causation is sharply distinct from Hume's reductive skeptical event-based theory, Watkins thinks that we should then seriously reconsider Kant's rationalist metaphysics of causality, especially including his appeals to (i) intrinsic causal powers, (ii) the necessitarian conception of causal laws, (iii) substance-based causation, and in particular, (iv) agent causation:

[J]ust as natural causality is to be understood primarily not in terms of events, but rather in terms of a substance determining the state of another substance by means of an exercise of causal powers in accordance with its nature, so, too, freedom is to be understood not primarily in terms of desires (which are simply one kind of mental event), but rather in terms of an agent or, more metaphysically, a substance determining its actions according to its freely chosen character. (KMC, 14)

I should admit right up front that my own perspective on this set of issues strongly favors Watkin's line of argument, but with three additional twists. I agree with him that the reductive skeptical Humean analysis of causation, considered as an interpretative backdrop, is insufficient for understanding the core notions of Kant's metaphysics of causality. And I also agree with him that Kant's metaphysics of causality in both the pre-Critical and Critical periods is at least as much rationalist-driven as it is empiricist-driven. But (1) I think that the Second Analogy and Third Antinomy are best read as attempts to *fuse* the rationalist and empiricist accounts of causation in a distinctively non-rationalist and non-empiricist way within the framework of transcendental idealism, and furthermore to contribute directly to a radically new Kantian philosophy of nature that begins to emerge explicitly in the *Metaphysical Foundations of Natural Science*. Also (2) I think that Kant's metaphysics of causality in the *post*-Critical period after 1787 — especially including the *Critique of Practical Reason*, the *Critique of the Power of Judgment*, and the *Opus postumum* — introduces several original metaphysical factors unrecognized by either the rationalist and empiricist approaches, that belong to this radically new Kantian philosophy of nature. Then finally (3), as a consequence of (1) and (2), I think that Kant's post-Critical metaphysics of causality and his new philosophy of nature

together yield the outlines of an adequate solution to the problem of causation. As I have said, these points are properly speaking only riffs or spins on Watkins's account, not disagreements. He explicitly says in a footnote that

Following [my] approach to its logical conclusion, one ought also to consider Kant's views on teleological causation in the *Critique of Judgment* as well as his views on physics in the *Metaphysical Foundations of Nature Science*. (KMC, 7n.)

Of course you cannot do everything in a single book! So my critical remarks in the final section are intended only to provide some suggestions about how one might go about developing the important and rich line of thought that Watkins has opened up.

III The Pre-Critical Kant and the Problem of Causation

All Kantians know that in the late 1760s the memory of his having read Hume's *Enquiry concerning Human Understanding* in the mid-1750s awoke Kant from a 'dogmatic slumber' that was filled with Leibnizian dreams, and propelled him towards the Critical Philosophy (P 4, 260). But fewer know that surprisingly many of those Leibnizian dreams resurfaced in waking life during Kant's Critical period as intrinsic parts of his transcendental idealism.

In chapter 1, 'Pre-Established Harmony versus Physical Influx,' Watkins embeds Kant's pre-Critical theory of causality in the 1740s, 50s, and 60s in the context of a vigorous and exclusively German philosophical conversation between Christian Wolff, Martin Knutzen, Alexander Baumgarten, G.F. Meier, and C.A. Crusius. One set of interlocutors — Wolff, Knutzen, Baumgarten, and Meier — were all Leibnizians of one stripe or another, and thus proponents of the Enlightenment. Crusius, by sharp contrast, was an anti-Leibnizian and a Pietist. So like all serious philosophical conversations, it also had a deeper moral source and a wider cultural resonance.

The purely philosophical part of the debate, however, centered on the Leibnizian question of whether the world operates according to the principle of pre-established harmony (which says that there exists a causal parallelism between finite substances or monads, created and determined by God) or according to the principle of physical influx (which says that there is real causal interaction between finite substances). Leibniz himself, as a theistic panexperientialist, defended both the principle of pre-established harmony and the idealistic thesis that all finite substances have intrinsic non-relational mental representational

properties upon which all their extrinsic spatial, temporal, and physical relations phenomenally strongly supervene.⁵ At the same time he rejected both physical influx and occasionalism (i.e., the thesis that all finite minds and material substances are created or recreated at every moment in time by God):

Leibniz argued that physical influx encountered problems explaining (1) how the mind and body could interact given their lack of homogeneity, (2) how it did not violate relevant conservation laws, and (3) how any two finite substances could act on each other if it was agreed that their accidents could not migrate from one substance to another. He argued that occasionalism was ultimately in no better shape, insofar as (1) it could not provide identity conditions for substances (given that God's action leaves no trace in things and there are no activities in substances that could explain how they could exist both as distinct from God and as numerically identical over time), (2) it was incompatible with human freedom (which Leibniz thought presupposed the activity or spontaneity of substances), and (3) it ran the danger of being committed to perpetual miracles by invoking God as the cause of everything. (KMC, 94)

Wolff accepted both Leibniz's idealism and his doctrine of pre-established harmony, but restricted its scope to the mind-body relation, and rather tantalizingly left open the question of whether some simple finite substances might enter directly into real causal relations without strongly supervening on intrinsic non-relational representational properties. Wolff never went so far as to assert the metaphysical possibility

5 More technically: *A*-facts (or higher-level facts) about the instantiation of *A*-properties strongly supervene on *B*-facts (or lower-level facts) about the instantiation of *B*-properties if and only if (1) necessarily if anything has a *B*-property, then it also has an *A*-property ('upwards determination'), and (2) necessarily there can be no change in anything's *A*-properties without a corresponding change in its *B*-properties ('necessary covariation'). Strong supervenience, which is a cross-possible-world modal dependency relation, can be distinguished from weak supervenience, which is a one-world relation. Also sometimes supervenience is defined strictly in terms of cross-world necessary co-variation without upwards dependency, which might be called 'moderate supervenience.' Still other types of supervenience track differences in the type of necessity (logical supervenience vs. natural or nomological or physical supervenience) or in the scope of the supervenience base (local supervenience vs. regional supervenience vs. global supervenience). A standard formulation of minimal materialism calls for the token-identity of all particular things or events of any kind with particular physical things or events ('token physicalism'), together with the strong logical global supervenience of all properties and facts on fundamental physical properties and facts. For more details, see J. Kim, *Supervenience and Mind* (Cambridge: Cambridge University Press 1993); and J. Kim, *Philosophy of Mind* (Boulder, CO: Westview Press 1996), ch. 1.

of physical influx; but on the other hand he did not definitively rule it out either.

Knutzen walked right through the door left slightly ajar by Wolff. Starting with Leibnizian assumptions to the effect that the simple substances constituting bodies have the power to move themselves and are impenetrable, he then explicitly defended physical influx and rejected pre-established harmony. This was possible because of an ambiguity in Leibniz's distinction between the primitive forces of monads and the derivative forces acting on physical bodies. The derivative forces causally strongly supervene on the monadic primitive forces, yet the derivative forces of physical bodies also may seem to have 'downward' causal influence on the primitive forces by being able to act on themselves or on their own bodies. In fact, for Leibniz the causal powers of the derivative forces would be excluded or trumped by the causal powers of the primitive forces and are strictly speaking epiphenomenal, although by God's command they operate strictly in parallel with the primitive mental forces. So Knutzen quite wrongly, but also quite interestingly, inferred the existence of physical influx from the causal powers of the supervenient epiphenomenal derivative forces — thus, in effect, anticipating the basic troubles of contemporary non-reductive materialism by 250 years.⁶

Baumgarten and Meier replied to Knutzen, and also returned the discussion more or less to the orthodox Leibnizian fold, by not (as Wolff had done) restricting pre-established harmony to the mind-body relation, and by explicitly offering neo-Leibnizian arguments in favor of it:

Baumgarten's main argument for pre-established harmony turns on the idea that God would create the world that had the greatest amount of harmony among all its substances. Such a world, he thinks, would be governed by pre-established harmony, since according to pre-established harmony, every state of every substance harmonizes with everything else... Meier goes even farther than Baumgarten insofar as he constructs several arguments in support of pre-established harmony that are based on the 'mechanics' of action in ways that go beyond what Leibniz has to say. For his arguments rely on principles such as 'When a finite substance acts, its inner state is thereby changed,' 'a smallest effect must be possible,' and 'the consequences remain the same when the grounds are the same,' principles that Leibniz never explicitly endorses. (KMC, 98)

In response to these neo-Leibnizians, Crusius developed an importantly different view according to which the finite substances that make

6 See J. Kim, 'The Myth of Nonreductive Materialism,' and 'The Nonreductivist's Troubles with Mental Causation,' both in Kim, *Supervenience and Mind*, 265-84, and 336-57.

up the material world mutually depend on one another for their existence. This was based on his idea that there is a fundamental distinction between God's understanding and God's will, and correspondingly between the principle of sufficient reason and 'the principle of determining reason':

If, following Leibniz, every substance were a world apart from all others, then only the fact that God thinks of them as belonging to a single world makes them belong to the same world. Crusius provides a striking contrast to this picture by arguing that substances must stand in real causal connections in order to have the kind of unity that is required to form a single world. And the ground for this connection cannot lie exclusively in God's understanding, as Leibniz would have it, but must rather depend, at least in part, on God's will. Moreover, Crusius thinks that substances belonging to a single world must be connected either by active grounds or by what he introduces under the term 'existential grounds.' While existential grounds can be understood in terms of the principle of sufficient reason and the principle of contradiction, active grounds are based on principles that extend beyond the principle of contradiction. In particular, Crusius distinguishes between the principle of sufficient reason and the principle of determining reason in order to allow for a distinction between libertarian free actions and events that are determined in the ordinary course of nature. (KMC, 98)

The Crusius-inspired idea that the ground of physical influx is a world made up of existentially interdependent individual substances then became the basis of Kant's early theory of causation, which Watkins works out in chapter 2, 'Kant's Pre-Critical Theory of Causality.' Kant, like Crusius, rejected pre-established harmony:

If the central question of the mid-1740s and for some time thereafter in German was whether to accept pre-established harmony or physical influx, Kant comes down decisively against pre-established harmony in favor of physical influx. Moreover, Kant arrives at his position by assuming (with Wolff and his followers) that monads are not necessarily mental, but could be physical instead. That is, Kant holds that what is substantial and real are principles of unity that are located in space and are the seats of physical forces. Kant, like many others of the period, then attempts to provide a metaphysically satisfying explanation of the physical properties of bodies, ending up with the commonsensical view that substances (i.e., physical monads) interact with each other causally so as to bring about the physical properties of bodies.... What is distinctive about Kant's *True Estimation* and what represents its real contribution is not what he had hoped for, namely a resolution of the *vis viva* [living force — R.H.] controversy, but rather his idea that a force, which brings about specific properties of bodies such as motion and impenetrability, should be understood not in terms of its specific effects, such as *vis motrix* or moving force (which Wolffians had done), but rather as essentially active. (KMC, 177-8)

A crucial difference between Kant and Crusius, however, was that for Kant finite physical substances do not depend on one another merely existentially but rather are interdependent by virtue of a universal system of essentially active forces coordinated by God:

In the *Nova dilucidatio's* principle of coexistence, Kant turns to explaining how the kind of mutual interaction between substances that is necessary for change is at all possible. Though Kant's argument is complex and subtle, his main idea is that God alone can coordinate the causal interaction between substances such that harmonious states result. This principle represents a significant departure from Crusius's position, because Crusius thought that substances could stand in real relations by means of their existence alone. Kant argues, however, that substances could exist and either not be related at all or be related ('placed' or 'situated') in a variety of ways. Kant infers from this that there must be a ground that determines which relational determinations hold between substances, just as there must be grounds for intrinsic determinations between substances. (KMC, 178-9)

In my opinion, the most important feature of Kant's pre-Critical theory of causality is his idea that the interactive forces between physical substances are expressed by what one can call (and this is my term, not Watkins's) the *intrinsic relational causal powers* of the individual material substances, and that these powers are not mere extrinsic dispositions of those material substances:

Even if it is clear that the finite substances that stand in mutual interaction are not in a position to make that interaction possible, why go so far as to think that what makes the interaction possible must be the cause of the very existence of the substance as well? Kant's answer depends on the fact that the relational properties of substances concern the very structure of a substance and therefore that whatever causes them must also be the cause of the existence of the substance insofar as what causes the existence of the substance would also cause its basic structure. In what sense do relational properties concern the very structure of a substance? As we saw, not only do relational properties go beyond intrinsic [non-relational — R.H.] properties, but it is also the case that whether a substance will have relational properties at all (and if so, which ones it will have) is a fundamental question, affecting what a substance can cause in other substances, just as its intrinsic [non-relational — R.H.] properties do. Thus in this sense, too, the relational properties of substances are just as fundamental as their intrinsic [non-relational — R.H.] properties. (KMC, 152)

In *The True Estimation of Living Forces* (1747) Kant held that these intrinsic relational causal powers express self-generated or spontaneous active 'living forces,' as opposed to externally-imposed or inertial active 'dead forces' (KMC, 108). But later in the *Nova dilucidatio* (1755) and the *Physical Monadology* (1756), he adopted the Newtonian mechanistic idea that the causal powers are not self-generated or spontaneous but instead inertial, as the result of more fundamental universal attractive and repulsive active forces operating on all material substances.

The other important causal notion that Kant developed in the *Nova dilucidatio* and the *Physical Monadology* was the result of an initial response to Hume's first *Enquiry*. Hume had argued that since (i) causal necessity is the same as logical sufficiency, but (ii) all things or events in nature are logically independent of one another, therefore (iii) there can

be no causal necessity in nature. Kant responded — in a way highly reminiscent of Crusius's distinction between the principle of sufficient reason and the principle of determining reason — by distinguishing between two different types of sufficient grounds of necessitation, namely *logical* grounds and *real* grounds. Real grounds yield a necessitation that is not logically necessary. Causal necessity in nature can therefore be the result of real grounds, even if all natural things and events are logically independent of one another. This, I think, is the first anticipation of what I call Kant's *modal dualism*, the thesis that there are two irreducibly different types of necessity: (1) logical or analytic necessity, and (2) non-logical or synthetic necessity.⁷

IV The Critical Kant and the Problem of Causation

Kant's 'Critical turn' in the late 1760s and early 70s seems to have been the joint result of remembering Hume's attack on the rationalist idea of causal necessity as a logical sufficiency relation between things or events (*P* 4: 260-261), together with a revolutionary insight achieved in 1769: 'the year '69 gave me great light' (*R* 5037; 18: 69). In turn, Kant's revolutionary insight was composed of two conjoined thoughts. His first thought was that the fundamental task of metaphysics is neither to investigate the ontology of things-in-themselves (which are unknowable) nor to investigate the psychologistic epistemology of sensory knowledge (which fails to yield either certainty or necessity), but rather to understand how mental representations — especially those that are a priori — can refer to their objects and be meaningful (I will call this doctrine Kant's *cognitive semantics*) (*PC* 10: 129-30). And his second thought was that human sensibility has two a priori subjective forms, the representations of space and time, that not only make a direct and necessary semantic contribution to all meaningful cognition, but are also

7 I should emphasize that the thesis that Kant's analytic-synthetic distinction entails modal dualism is mine, not Watkins's. Moreover in my opinion, at this point in Kant's philosophical development, since he had not yet recognized that the fundamental semantic distinction between intuitional representational content and conceptual representational content is the basis of the distinction between analytic judgments and synthetic judgments, he was unable to assert the irreducibility of real grounds to logical grounds. This is because, without the sharp intuition vs. concept distinction, real grounds can still be construed as merely *conditional* logical grounds. That is, non-logical necessitation can still be construed as nothing but logical entailment under an antecedent material condition. See Hanna, *Kant and the Foundations of Analytic Philosophy*, ch. 5.

irreducible to concepts, judgments, and discursive or logical representations more generally, and above all are identical to space and time themselves (Kant later calls this doctrine *the transcendental aesthetic*) (ID 2: 398-406). Taking Hume's attack on rationalism seriously, together with his cognitive semantics and the transcendental aesthetic, then yields Kant's characteristic thesis that irreducibly non-logical or synthetic a priori necessitation is the basis of all causal relations, and thereby correspondingly also yields the modal dualism that is in my opinion essential to the transcendental metaphysics of his Critical period. This of course he famously describes in the first *Critique* as the philosophical project of answering the question '**how are synthetic a priori judgments possible?**' (CPR B19).

Watkins's way of formulating the Critical turn is slightly different, although I think quite compatible with this picture. For Watkins, Kant's basic moves are these: We cannot know things-in-themselves through our finite human senses. But we can and do know appearances both through our senses and through our understanding, by means of objectively valid empirical judgments. This knowledge in turn has necessary a priori metaphysical conditions, expressible as strict normative Principles of Pure Understanding, which turn out to impose a set of broadly rationalistic formal constraints on the existence, intrinsic properties, and interrelations of all appearances (KMC, 181-8).

But however we formulate the Critical turn, one direct result is that according to Kant, causation as analyzed in the Second Analogy turns out to inhere neither in things-in-themselves nor in the contingent succession of sensory events, but rather instead in the formal a priori necessary conditions for representing and knowing how the successive states of apparent objects are strictly determined under natural laws in asymmetric time. This is also known in contemporary Kant-literature as 'the problem of time determination.' Watkins develops his interpretation of Kant's solution to the problem of time determination in chapter 3, 'Kant's Second and Third Analogies of Experience.' The leading idea of Watkins's account is that the Third Analogy is as essential to solving the problem of time determination as the Second Analogy.

A basic presupposition of Kant's solution for the problem of time-determination is his pre-Critical idea, borrowed from Newtonian mechanism, that no individual material substance determines changes in itself but on the contrary all its changes must be imposed on it externally by something else. Therefore it seems that either (1) the *earlier* states of objects strictly nomologically determine the *later* states of the same or different objects, by virtue of their extrinsic relations of irreversible temporal ordering (event-causation), or else (2) *objects alone* nomologically strictly determine the states of different objects, by virtue of their intrinsic causal powers (substance-causation). The standard Hume-ori-

ented view of Kant's analysis in the Second Analogy strongly favors the first option. But as Watkins correctly points out, this then makes it impossible to understand the universal mutual interaction of objects described in the Third Analogy, which requires a world of *simultaneous substances*. So he concludes that we must adopt the second or rationalist option as the most charitable and philosophically coherent interpretation of the Second and Third Analogies taken together:

Kant's Second and Third Analogies of Experience attempt to argue that causality and mutual interaction (metaphysical relations) are necessary conditions for knowledge of objective succession and coexistence (epistemological items), which are in turn required for several different unities: the unity of nature/world, the unity of time, the unity of experience, and the unity of apperception. We thus now have confirmation of the speculative thesis advanced at the beginning of the chapter that the nature of Kant's 'Critical turn' ought to consist in a mixture of continuities and breaks compared with his pre-Critical views. The discontinuities are most apparent (1) in the addition of a series of epistemological concepts (e.g., the unities of apperception and experience) to the argumentative framework Kant employs (such that his primary concern is with, e.g., phenomenal rather than noumenal substances), and (2) in the fact that causality and mutual interaction are alleged to be necessary for *knowledge* of succession and coexistence. It is thus in line with his metaphilosophical rejection of mere analysis as a legitimate means of providing knowledge of synthetic a priori truths that Kant employs experience (of succession and coexistence) in establishing metaphysical principles (of causation and mutual interaction). However, what has not yet been noted in any detailed way by commentators and what is thus of special significance for our understanding of Kant's views are the widespread continuities. The most obvious one pertains to the basic content of Kant's claims in the *Nova dilucidatio's* principles of succession and coexistence and in the *Critique's* Second and Third Analogies, namely that causal relations are required for the temporal relations of succession and coexistence. Another continuity, much less apparent to those who have focused almost exclusively on the Second Analogy, is Kant's assumption that a substance cannot act on itself so as to change itself or to determine its place in time, an assumption Kant makes most clearly in the *Nova dilucidatio's* principle of succession. (KMC, 229)

This is not, however, to say that Watkins has simply re-packaged the famous Patchwork Thesis developed by Hans Vaihinger and Norman Kemp Smith, which says that Kant incoherently runs together pre-Critical and Critical ideas in the first *Critique*.⁸ On the contrary, Watkins is saying that *the full and proper sense* of Kant's Critical turn in the *Critique of Pure Reason* can be understood *only* by seeing how several of its basic ideas have *essentially* pre-Critical foundations.

8 See, e.g., F. Rauscher, 'Review of *Kant and the Metaphysics of Causality*,' *Notre Dame Philosophical Reviews* (November 2005), URL= <<http://ndpr.nd.edu/cfm?id=4521>>.

This idea is further developed in chapter 4, 'Kant's Model of Causality.' On Watkins's reading of the Second and Third Analogies, Kant's Critical theory of causation is grounded on his pre-Critical idea of what I called the 'intrinsic relational causal powers' of substances, whose realizations in the physical world are essentially expressions of the universal active inertial forces of attraction and repulsion. This is what Kant calls 'the causality of the cause.' The basic thought behind this notion is that since no substance can determine its own states, then every state of a substance is caused by some antecedently or simultaneously existing substance that realizes one or more of its causal powers by making another substance determinate in some respect or another. The activity by which a substance makes another substance determinate is itself inherently *indeterminate*, and in that way the indeterminate activity of individual substances is the causally determining ground of the determinate nature of the actual world. This view entails a rejection of event-based causation in favor of substance-based causation, and is rationally forced on us by taking the Third Analogy seriously:

The case against event-event causation is made by showing that an event-event model is inconsistent with Kant's claim in the Third Analogy that mutual interaction is necessary for knowledge of coexistence. For it is simply incoherent (in the context of Kant's philosophy) to assert that two events could stand in mutual interaction with each other. (KMC, 231)

By attending to Kant's notion of mutual interaction in the Third Analogy and to various nonargumentative passages in the Second Analogy, we have seen that Kant neither does nor can accept Hume's event-event model of causality. Rather, with the benefit of an awareness of Kant's pre-Critical account of causality, we saw that Kant's texts and arguments commit him to a model of causality that involves substances exercising their causal powers so as to determine each other's states. While accepting causal powers is quite traditional in the context of early modern philosophy, Kant develops his notion of causal powers in detail and incorporates it into his broader account of (phenomenal) metaphysics. More specifically, the exercise of a causal power or, as Kant sometimes puts it, 'the causality of the cause,' is not a determinate event but rather an asymmetrical and indeterminate activity that brings about passive determinations in a distinct substance. (KMC, 296)

Watkins's reading of Kant's theory of causation also has important implications for the vigorous recent debate about the nature of causal laws in Kant's metaphysics.⁹ For Watkins, the causal laws are synthetic

9 The problem is how to understand both the apparently a priori epistemological and also strongly modal status of these laws, in view of the fact that they are explicitly held to be empirical. See, e.g., H. Allison, 'Causality and Causal Laws in Kant: A Critique of Michael Friedman,' in P. Parrini, ed., *Kant and Contemporary Epistemology*

a priori propositional rules that strongly supervene on what I called the intrinsic relational causal powers inhering in individual material substances, and strictly govern the production of changes in the states of different material substances. Causal laws are therefore both non-logically necessary and also explanatorily and ontologically dependent on the natures of individual substances (KMC, 288-9).

While I endorse Watkins's rejection of the simple Humean analysis of the Second Analogy and also his combinatory reading of the Second and Third Analogies, and while I also endorse the idea that for Kant causal laws are necessarily connected to the intrinsic relational properties of individual substances, I also think that there is a third alternative for interpreting the Second and Third Analogies. Suppose that individual material substances are *themselves* strongly supervenient on sets of primitive strict determination-relations running between earlier and later states of the same underlying world-totality of matter, as described in the First Analogy, which in turn according to Kant in the *Metaphysical Foundations of Natural Science* is nothing but a dynamic totality of attractive and repulsive forces (MFNS 4, 534). The strong supervenience of individual substances on the attractive and repulsive forces allows for the emergence of novel irreducible relational properties at the higher level of those substances. Then individual substances can be at once 'upwardly' strictly determined by their asymmetric successive event-relations (diachronic supervenience), and also stand in 'sideways' or simultaneous dynamic relations of mutual interaction (synchronic mutual determination). On this alternative Kantian picture, synthetic a priori causal laws would then be irreducible emergent *second-order* intrinsic relational properties of individual material substances, grounded on the primitive active inertial dynamics of the universal attractive and repulsive forces, embedded within an overarching framework of transcenden-

(Netherlands: Kluwer 1994) 291-307; G. Buchdahl, *Metaphysics and the Philosophy of Science* (Cambridge: The MIT Press 1969) 651-65; G. Buchdahl, 'The Conception of Lawlikeness in Kant's Philosophy of Science,' in L.W. Beck, ed., *Kant's Theory of Knowledge* (Dordrecht: D. Reidel 1974), 128-50; P. Guyer, 'Kant's Conception of Empirical Law,' in P. Guyer, *Kant's System of Nature and Freedom* (Oxford: Oxford University Press 2005), ch. 2; M. Friedman, *Kant and the Exact Sciences* (Cambridge: Harvard University Press 1992), chs. 3-4; M. Friedman, 'Causal Laws and the Foundations of Natural Science,' in P. Guyer, ed., *The Cambridge Companion to Kant* (Cambridge: Cambridge University Press 1992), 161-99; W. Harper, 'Kant on the A Priori and Material Necessity,' in R. Butts, ed., *Kant's Philosophy of Physical Science* (Dordrecht: D. Reidel 1986), 239-72; R. Walker, 'Kant's Conception of Empirical Law,' *Proceedings of the Aristotelian Society*, Supp. Vol. 63 (1990) 243-58; and E. Watkins, 'Kant's Justification of the Laws of Mechanics,' in E. Watkins, ed., *Kant and the Sciences* (New York: Oxford University Press 2001) 136-59.

tal laws expressed by the System of Principles. And this in turn would allow Kant to combine event-causation and substance-causation in a single elegant metaphysical scheme, and thereby fuse the empiricist-Humean and rationalist-Leibnizian approaches to the metaphysics of causality in a way that is importantly different from either of the classical approaches. Not only that, it also makes good interpretive sense of the otherwise puzzling fact that Kant switches freely back and forth in the three Analogies between event-causation talk and substance-causation talk.

On the other hand however, a crucial double pay-off of Watkins's more rationalist-oriented interpretation of Kant's analysis of causality in the Second and Third Analogies is that it not only very effectively accounts for Kant's transcendental resolution of the Third Antinomy, but also has important implications for our contemporary debate about the problem of free will.¹⁰ This thought is carefully worked out in chapters 5 and 6, 'The Metaphysics of Freedom' and 'Kant's Reply to Hume: Historical and Contemporary Considerations.'

The Third Antinomy describes the seemingly unresolvable metaphysical opposition between universal natural determinism and the existence of spontaneous non-deterministic free will. Kant argues that the metaphysical opposition is logically fallacious because it fails to heed the transcendental idealist distinction between (a) appearances or phenomena, and (b) non-sensible entities or noumena. Granting the phenomena-noumena distinction however, universal natural determinism and free will are then mutually consistent if phenomena are naturally determined and noumena are free.

But precisely what sort of causation is *noumenal* causation or 'transcendental freedom'? Kant says this:

By freedom in the cosmological sense ... I understand the faculty of beginning a state **from itself** (*von selbst*), the causality of which does not in turn stand under another cause determining it in time in accordance with the law of nature. Freedom in this signification is a pure transcendental idea, which, first, contains nothing borrowed from experience, and second, the object of which also cannot be given determinately in any experience, because it is a universal law — even of the possibility of experience — that everything that happens must have a cause, and hence that the causality of the cause, as **itself having happened** or arisen, must in turn have a cause; through this law, then, the entire field of experience, however far it may reach, is transformed into the sum total of mere nature. But since in such a way no absolute totality of conditions in causal relations is forthcoming, reason creates the idea of a

10 See, e.g., R. Kane, *A Contemporary Introduction to Free Will* (Oxford: Oxford University Press 2005); R. Kane, ed., *The Oxford Handbook of Free Will* (Oxford: Oxford University Press 2002); and Watson, ed., *Free Will*.

spontaneity, which could start to act from itself, without needing to be preceded by any other cause that in turn determines it to action according to the law of causal connection. It is especially noteworthy that it is this **transcendental** idea of **freedom** on which the practical concept is grounded.... **Freedom in the practical sense** is the independence of the power of choice (*Willkür*) from **necessitation** by impulses of sensibility. (CPR A533/B561)

Transcendental freedom is absolutely spontaneous causation: causation that is unprecedented or original, productive, and self-guiding, and has no nomologically sufficient temporally antecedent conditions. If I have transcendental freedom of the will, then I am the ultimate source of at least some of my choices and actions. So in a deep or metaphysically real sense, those choices and actions are *up to me*. Furthermore transcendental freedom is a necessary but not sufficient condition of *practical freedom or autonomy*, that is, free intentional action that (in a negative sense) is undetermined by all alien causes especially including egoistic sensible impulses, and (in a positive sense) consists in self-legislation according to the universal moral law or Categorical Imperative (GMM 4, 446-7). If Watkins is correct, then Kant's theory of noumenal causality directly entails that a transcendently free cause is an *agent cause*, namely, an immaterial, non-spatiotemporal substance that stands outside the naturally-determined series of phenomenal events and absolutely spontaneously determines changes in material substances.

By deploying his account of Kant's model of causality, Watkins is able to offer a Kantian interpretation of agent causation that gets around a deep and notorious problem for classical Humean compatibilist or New Compatibilist (such as Harry Frankfurt's) desire-based and event-based theories of intentional action. The problem is this. Compatibilists assert that free will and determinism are actually and metaphysically consistent. But since according to determinism every event, whether it is a physical event like an intentional body movement or a psychological event like a conscious desire or intention, is itself nomologically sufficiently caused by some antecedent event, and since the regress of such events runs back at least to the Big Bang, it follows that none of my intentional movements or choices is ever up to me. So Humeans and New Compatibilists cannot explain either our causal control over our choices and doings, or our moral responsibility for them, and free will is thereby rendered impossible. And in this way compatibilism collapses into 'hard determinism,' which says that determinism exists and free will is metaphysically impossible. But even if Hume and the New Compatibilists can't do the job, Kant still can:

By drawing on features of Kant's general model of causality, especially the notion of determination that it employs, we see how he can develop plausible responses to problems that arise in the context of free will and determinism. The regress that

seems to arise in explaining a free action in terms of desires can be stopped (and the regress problem solved) because Kant's general model of causality employs the notion of something other than a desire, namely that of a determining ground, which cannot be explained by anything external to itself, since qua determining ground it is not determinable, that is, cannot be determined by anything other than itself. The location problem can likewise be solved by appealing to a notion employed in Kant's general model of causality other than that of an event, namely the notion of a substance that determines its actions. For insofar as an agent (or self) that freely determines its actions can be a substance, it is not an event that could be determined by anything else and thus does not run the risk of being determined by something else in such a way that it must abdicate responsibility for its actions. Accordingly, Kant's notion of a determination can not only be employed in both his account of freedom and his general model of causality, but can also be used to solve a series of traditional problems and explain a variety of phenomena that are not easily incorporated into a single coherent philosophical account. (KMC, 360)

This Kantian doctrine of noumenal agent causation, which is also known as 'the timeless agency theory,'¹¹ in turn directly corresponds to the libertarian indeterminist incompatibilist position on free will iconoclastically defended in the second half of the 20th century by Roderick Chisholm¹² in the face of an almost universal acceptance of either hard determinist incompatibilism on the one hand, or 'soft' determinist Hume-style or Frankfurt-style compatibilism on the other. And in this way, Kant's non-Humean, rationalist, transcendental analysis of causation and his corresponding libertarian indeterminist incompatibilist analysis of free will can be used to buttress a fruitful trend in the early 21st century towards the serious reconsideration of agent-causation.¹³

Looked at from the standpoint of the Third Antinomy and its transcendental resolution, the elegance and power of Watkins's interpretation of Kant's metaphysics of causality cannot be denied. But it seems to me that there are at least three deep philosophical problems with the very idea of timeless agent causation (or TAC for short). First, TAC is directly in conflict with our immediate consciousness of our own intentional agency, which presents us as ineluctably embodied, conscious desiring practically rational animals in time and space. Second, TAC is seemingly committed to interactionist substance dualism, which in turn directly

11 See A. Wood, 'Kant's Compatibilism,' in A. Wood, ed., *Self and Nature in Kant's Philosophy* (Ithaca, NY: Cornell University Press 1984) 73-101; and D. Pereboom, 'Kant on Transcendental Freedom,' *Philosophy and Phenomenological Research*, forthcoming.

12 R. Chisholm, 'Human Freedom and the Self,' in Watson, ed., *Free Will*, 26-37

13 See, e.g., T. O'Connor, *Persons and Causes: The Metaphysics of Free Will* (New York: Oxford University Press 2000).

violates the highly plausible Principle of the Causal Closure of the Physical, to the effect that only physical things can cause physical things. And third, TAC is most certainly committed to causal overdetermination, which directly violates the equally highly plausible Explanatory Exclusion Principle, to the effect that there can be at most one complete and independent cause of any event.¹⁴ So if Kant really is a TAC theorist, then I think that he is in Really Big Trouble.

The other standard approach to Kant's theory of freedom in the contemporary Kant-literature is known as 'the two standpoints theory.'¹⁵ This theory, which basically follows the line of argument developed by Kant in the third chapter of *Groundwork of the Metaphysics of Morals*, says two things. First, persons can act only under the pure rational idea of their own practical freedom or autonomy, lest they pragmatically contradict their own rational agency and undermine the possibility of their moral responsibility. Second, persons can act under the pure rational idea of their own freedom only by regarding themselves 'from the noumenal standpoint' as causally efficacious things-in-themselves, hence as transcendently free causes, even though absolutely spontaneous transcendental freedom is itself scientifically inexplicable. Ironically, the main theoretical vice of the two standpoints theory is what is also generally taken to be its main theoretical virtue, namely that it is compatibilist. But since the two standpoints theory yields only a *belief* in our freedom, then if, as per compatibilism, we really *are* also naturally determined, and also if, as per fallibilism, our belief in our own free will also actually is *false*, then we are nothing but naturally mechanized puppets epiphenomenally dreaming we are real persons. And that is not my idea of a good time. So if Kant really is a two standpoint theorist, then I think that again he is in Really Big Trouble.

So what is a contemporary Kantian to do? Fortunately there is, I think, a third and philosophically much more defensible metaphysics of causation available in Kant's writings. According to this third theory, transcendently free will is a non-mechanistic yet still fully natural process that (1) can occur in all and only rational desiring animals, and (2) possesses diachronically emergent, causally efficacious, and yet also irreducibly non-sensible or 'negatively noumenal' (CPR: B307) norma-

14 See J. Kim, 'Mechanism, Purpose, and Explanatory Exclusion,' in Kim, *Supervenience and Mind* 237-64.

15 See, e.g., H. Allison, *Kant's Theory of Freedom* (Cambridge: Cambridge University Press 1990).

tive and modal properties. Let us call this *the embodied agency theory*¹⁶ In order to find the embodied agency theory, I think that we must turn to the *post-Critical Kant*.

V The Post-Critical Kant and the Problem of Causation

By 'the post-Critical Kant' I mean the corpus of Kant's published or unpublished writings in metaphysics and ethics *after 1787*, the year of the publication of the B edition of the first *Critique*. Kant's post-Critical era includes the *Critique of Practical Reason* (1788), the *Critique of the Power of Judgment* (1790), and the drafts of the unfinished *Transition from the Metaphysical Foundations of Natural Science to Physics* written in the late 1790s, later published as the *Opus postumum*.

Essential to the possibility of recognizing a post-Critical Kant who is importantly distinct from both the pre-Critical Kant and the Critical Kant, is a preliminary reconsideration of his theory of physical matter in the pre-Critical period. As early as the *Nova dilucidatio* and the *Physical Monadology*, Kant developed the doctrine of what Thomas Holden aptly calls 'the Kant-Boscovich force-shell atom theory,' according to which matter is composed of indivisible extensionless points that radiate surrounding infinitely divisible shells of inertial active force.¹⁷ The force-shell atom theory anticipated Michael Faraday's field theory by 100 years. But above all it demonstrates Kant's deep deviance from classical views of the nature of matter. Like Boscovich, Kant thought of the basic elements of matter as essentially positions in a larger causal-dynamic structure, not as primitive static atoms defined by their intrinsic non-relational properties. Otherwise put, in the Kant-Boscovich theory of matter *causal-dynamic structure and function are explanatorily and ontologically prior to material substance*. In the 19th century Nietzsche presciently grasped the truly radical nature of the Kant-Boscovich structuralist metaphysics of matter:

While Copernicus has persuaded us to believe, contrary to all the senses, that the earth does *not* remain fixed, Boscovich has taught us to give up the belief in the final

16 See R. Hanna, *Kant, Science, and Human Nature* (Oxford: Oxford Univ. Press, forthcoming), ch. 8.

17 T. Holden, *The Architecture of Matter* (Oxford: Clarendon/OUP 2004), ch. 6. Roger Boscovich independently developed essentially the same idea from a different starting point, and published it in 1758 in his *Theory of Natural Philosophy*.

'fixed' thing about the earth — the belief in 'substance,' in 'matter,' in the little residual earthly clump — the atom. This was the greatest triumph over the senses ever achieved on earth.¹⁸

Fast forward now to the 1770s and 80s. What above all distinguishes the Critical Kant of that era from the pre-Critical Kant of the 1750s and 60s, is the metaphysics of transcendental idealism he worked out in the *Critique of Pure Reason* (1781/1787), the *Prolegomena to Any Future Metaphysics* (1783), and the *Metaphysical Foundations of Natural Science* (1786). Correspondingly then, there are two defining features of the *post*-Critical Kant, both of which have to do with *closing gaps* in his Critical metaphysics of transcendental idealism.

(1) First, Kant made a serious attempt to close two inferential gaps in the Transcendental Deduction of the Categories. The first inferential gap is between the general transcendental laws of nature, and the increasingly specific empirical laws of nature (I call this *the top-down problem*). The top-down problem, in a nutshell, is that the existence of increasingly specific empirical laws cannot be syllogistically deduced from the existence of the general transcendental laws. And the second inferential gap is between what is originally given nonconceptually to the mind via causal affection in empirical intuition together with the a priori forms of intuition,¹⁹ and what is imposed on perceptual inputs by the conceptual structures of empirical judgment together with the pure concepts of the understanding (I call this *the bottom-up problem*). The bottom-up problem, in a nutshell, is that even granting the soundness of the B Deduction, it seems to be impossible to rule out altogether the possibility that nomologically ill-behaved natural objects falling outside the scope of the categories might still be given in nonconceptual intuition.

(2) Second, Kant also made a serious attempt to close an even more gaping explanatory and ontological gap between the deterministic mechanism of physical *nature* on the one hand, and the absolute spontaneity of transcendental and practical *freedom* on the other.

How did these attempted gap-closings go? Kant attempted to bridge the first gap and thereby solve both the top-down and bottom-up problems (which are in fact just mirror images of one another), by further developing his structuralist metaphysics of matter, and by appealing to teleological reasoning in the natural sciences. The earlier totality of

18 F. Nietzsche, *Beyond Good and Evil*, trans. W. Kaufman (New York: Random House 1966), sect. 12, 20, translation modified slightly.

19 See R. Hanna, 'Kant and Nonconceptual Content,' *European Journal of Philosophy* 13 (2005) 247-90.

force-shell atoms becomes a fundamental causal-dynamic fluid aether. This in turn operates as a global supervenience base for macrophysical material objects. Then the existence of this fundamental causal-dynamic fluid aether becomes a *transcendental material condition* for completing the Deduction of the Categories.²⁰ This in turn directly connects with Kant's striking idea, developed in the First Introduction and first half of the *Critique of the Power of Judgment*, that the experience of natural beauty presupposes a necessary a priori belief in an undesigned pre-adjustment or pre-harmonization of the lawful form of the material natural world to the operations of our basic cognitive faculties. In other words, Kant's idea is that in aesthetic experience of the beautiful in physical nature, nature must present itself to us as accidentally such that it can formally support our scientific inferences about it (this is what he calls 'purposiveness without a purpose'): so physical nature is accidentally but also inherently *user-friendly*. Lucky us.

Kant attempted to bridge the even bigger second gap in his Critical system in two ways. First, he appealed to the 'fact of reason' (*Faktum der Vernunft*) (CPrR 5: 31) and the inherently altruistic moral feeling of 'respect' (*Achtung*) (CPrR 5: 72-89), together with his earlier thesis that duty is the necessity of an action done from respect for the moral law (GMM 4: 400). This allows him to say that the act of freely doing one's duty is essentially an act of recognizing the moral law or Categorical Imperative as an overriding non-instrumental reason for action. But this act both normatively and motivationally determines the 'power of choice' (*Willkür*) — the faculty of effective first-order desires, i.e., the desires that move us all the way to action — by causally triggering the innate higher-order emotional disposition of respect, or the desire to have non-self-interested effective first-order desires, which is then implemented by the faculty of second-order volitions — the faculty of deliberative non-instrumental or instrumental reasoning about effective first-order desires and their ends — or the 'will' (*Wille*).

Second, Kant asserted that living organisms have an irreducibly non-mechanical teleological causal structure:

It is quite certain that we can never adequately come to know the organized beings and their internal possibility in accordance with merely mechanical principles of nature, let alone explain them; and indeed this is so certain that we can boldly say that it would be absurd for humans even to make such an attempt or to hope that

20 See, e.g., P. Guyer, 'Kant's Ether Deduction and the Possibility of Experience,' in Guyer, *Kant's System of Nature and Freedom*, ch. 3; and B. Hall, 'A Reconstruction of Kant's Ether Deduction in *Übergang* 11,' *British Journal of the History of Philosophy* (forthcoming).

there might yet arise a Newton who could make comprehensible even the generation of a single blade of grass according to natural laws that no intention has ordered; rather we must absolutely deny this insight to human beings. (CPJ 5: 400)

And he also put forward the striking metaphysical thesis that mind — and in particular the conscious desiring mind — entails biological life:

Life is the faculty of a being by which it acts according to the laws of the faculty of desire. (CPrR 5: 9n.)

Life without the feeling of the corporeal organ is merely consciousness of one's existence, but not a feeling of well- or ill-being, i.e., the promotion or inhibition of the powers of life; because the mind for itself is entirely life (the principle of life itself), and hindrances and promotions must be sought outside it, though in the human being himself, hence in combination with his body. (CPJ 5: 278, underlining added)

Now if biological properties are irreducibly non-mechanical and the mind is biologically alive, then the mind cannot be naturally mechanistically determined. This consequence, in turn, conforms smoothly to Kant's equally striking thesis in the Introduction to the *Metaphysical Foundations of Natural Science* that there can never be a natural science of empirical psychology:

The empirical doctrine of the soul must always remain ... removed ... from the rank of what may be called a natural science proper. This is because mathematics is inapplicable to the phenomena of the inner sense and their laws.... It can, therefore, never become anything more than a historical (and, as such, as much as possible) systematic natural doctrine of the inner sense, i.e., a natural description of the soul, but not a science of the soul. (MFNS 4: 471, underlining added).

If the conscious mind is biologically alive, not naturally mechanistically determined, and cannot be adequately described by natural science as Kant knew it, then either it operates under indeterministic strict statistical natural laws, or else its operations are merely random, or else it belongs to an embodied rational agent operating under non-mechanistic teleological and moral laws. The post-Critical Kant strongly believes that it is the latter.

Whatever else we may think of them, we must admit that Kant's post-Critical gap-closing ideas are all brilliantly original. They significantly influenced the post-Kantian idealist and romantic tradition — especially Herder, Schiller, Schelling, and Coleridge. And some of them have finally been rediscovered by contemporary philosophers. For example, Kant's dynamic structuralist conception of matter has been rediscovered by contemporary physicists under the rubric of field theory, and by contemporary metaphysicians under the rubric of 'atomless gunk'

theory.²¹ His distinction between the higher and lower faculties of desire, *Wille* and *Willkiir*, has been rediscovered under the rubric of the hierarchical desire model of the will.²² His theory of the rationally motivating force of the moral emotion of respect has been rediscovered under the rubric of conative objectivism.²³ His non-mechanistic theory of biological dynamics has been rediscovered under the rubric of dynamical systems theory.²⁴ And his idea that mind is identical to life has been rediscovered under the rubric of the strong continuity of mind and life thesis held by some biologically-oriented philosophers of mind:

Life and mind have a common abstract pattern or set of basic organizational properties. The ... properties characteristic of mind are an enriched version of the ... properties that are fundamental to life in general. Mind is literally *life-like*.²⁵

But most importantly of all, if Kant's brilliantly original post-Critical ideas really *were* true, and if we *were* willing to accept their radical philosophical consequences, then the very idea of causation would have to be re-thought, and the apparently insoluble problem of causation might then be finally solved. Or so it seems to me.

It must be emphasized that whether Kant's metaphysics of causality is best interpreted, as Watkins thinks, from the standpoint of his pre-Critical philosophy, as a *rationalist* non-empiricist theory, or instead best interpreted in the way I have suggested, from the standpoint of his post-Critical philosophy, as a radically original *non-rationalist and non-*

21 See J. Edwards, *Substance, Force, and the Possibility of Knowledge: On Kant's Philosophy of Material Nature* (Berkeley: University of California Press 2000); and D. Zimmerman, 'Could Extended Objects Be Made Out of Simple Parts? An Argument for "Atomless Gunk"', *Philosophy and Phenomenological Research* 56 (1996) 1-29.

22 H. Frankfurt, 'Freedom of the Will and the Concept of a Person,' in H. Frankfurt, *The Importance of What We Care About* (Cambridge: Cambridge University Press 1988) 11-25.

23 See A.W. Moore, *Noble in Reason, Infinite in Faculty: Themes and Variations in Kant's Moral and Religious Philosophy* (London: Routledge 2003); and R. Hanna and A.W. Moore, 'Reason, Freedom, and Kant: An Exchange,' *Kantian Review* (forthcoming).

24 See, e.g., J. Kelso, *Dynamic Patterns* (Cambridge: The MIT Press 1995); F. Varela, *Principles of Biological Autonomy* (New York: Elsevier/North-Holland 1979); and A. Weber and F. Varela, 'Life After Kant: Natural Purposes and the Autopoietic Foundations of Biological Individuality,' *Phenomenology and the Cognitive Sciences* 1 (2002) 97-125.

25 See P. Godfrey-Smith, *Complexity and the Function of Mind in Nature* (Cambridge: Cambridge University Press 1996), 320. See also E. Thompson, *Mind in Life* (Cambridge: Harvard University Press forthcoming).

empiricist theory, the considerable philosophical value of reading Kant's theory of causation as non-empiricist or non-Humean is nevertheless fully preserved. The novelist G.K. Chesterton, in his deliciously quirky book on the novels of Dickens, observed that although the eponymous hero of *Nicholas Nickleby*, Nicholas Nickleby, is a mere stick, *any* strong stick is good enough to beat Squeers. The same thing is true of the reductive skeptical Humean empiricist theory of causation. Any strong Kantian stick is good enough to beat it. The upshot is that Eric Watkins's *Kant's Metaphysics of Causality* is an absolutely first-rate book that every serious Kant scholar and everyone seriously interested in the problem of causation should study very carefully indeed.

Received April 2006

ROBERT HANNA
University of Colorado at Boulder
Boulder, CO 80309-0232
USA

