# Kant's Model of Causality: Causal Powers, Laws, and Kant's Reply to Hume

ERIC WATKINS*

KANT'S VIEWS ON CAUSALITY have received an extraordinary amount of scholarly attention, especially in comparison with Hume's position. For Hume presents powerful skeptical arguments concerning causality, yet Kant claims to have an adequate response to them. Since it turns out to be far from obvious what even the main lines of Kant's response are, scholars have felt the need to clarify his position. Unfortunately, however, no consensus has been reached about how best to understand Kant's views. To gauge the current state of the debate, consider briefly Hume's two-fold skeptical attack on causality in his *Enquiry Concerning Human Understanding* and what strategy Kant is often thought to be pursuing in response in the *Critique of Pure Reason*.[1]

[1] Hume's fuller views on causality are naturally more complex than the brief summary provided below indicates. For more comprehensive discussions, see, e.g., Barry Stroud, *Hume* (London: Routledge & Kegan Paul, 1978), John Wright, *The Sceptical Realism of David Hume* (Minneapolis: University of Minnesota Press, 1983), Richard Fogelin, *Hume's Skepticism in the Treatise of Human Nature* (Boston: Routledge & Kegan Paul, 1985), Galen Strawson, *The Secret Connexion: Causation, Realism, and David Hume* (New York: Oxford University Press, 1989), Ken Winkler, "The New Hume," *Philosophical Review* 100 (1991): 541–79, and Don Garrett "The Representation of Causation and Hume's Two Definitions of 'Cause'," *Nous* 27 (1993): 167–90. Moreover, the reception of Hume in Germany, documented by Günter Gawlick and Lothar Kreimendahl, *Hume in der deutschen Aufklärung: Umrisse einer Rezeptionsgeschichte* (Stuttgart-Bad Cannstatt: Frommann-Holzboog, 1987), and Manfred Kühn, *Scottish Common Sense in Germany, 1768–1800: A Contribution to the History of Critical Philosophy* (Kingston: McGill-Queen's University Press, 1987) only complicates matter further. While Hume's *Enquiry Concerning Human Understanding* was translated into German in 1755, only select quotations from Hume's *A Treatise of Human Nature* were published in German (in works by Beattie and Tetens and in Hamann's anonymous translation of I, 4, 7) before Kant wrote the *Critique of Pure Reason* in 1781. A further complicating factor is that Kant was, at times, in close contact with Hamann, who had translated, without publishing, some of Hume's texts concerning natural religion. Thus, it is possible that Hamann had translated, and that Kant had read, more of the *Treatise* than just I, 4, 7 and the selections included by Beattie and Tetens in their own works. Accordingly, Kant may (or may not) have read significant parts of Hume's *Treatise* by 1781. Moreover, since Hamann had certainly read the *Treatise,*

* **Eric Watkins** is Associate Professor of Philosophy at the University of California, San Diego.

*Journal of the History of Philosophy*, vol. 42, no. 4 (2004) 449–488

First, Hume argues that our idea of causality does not involve the idea of necessary connection or power that is traditionally[2] (or pre-theoretically) associated with it, since, according to Hume's copy principle, every idea must be copied from some impression, but nowhere do we have an impression of any such connection or power.[3] We have no *external* impression of a necessary connection between a cause and its effect, because all we ever perceive in the world is one event followed by another. What we see, for instance, in the collision of two billiard balls is the motion of one billiard ball followed by the motion of the other. We do not actually see, it is claimed, the one billiard ball *imparting motion* to the other. We have no *internal* impression of such a connection either, since we do not literally see (or even know at all) how the mind causes its own body to move or how it creates its ideas. In light of the fact that every cause seems to be completely distinct in our experience from its effect, we have no impression of any necessary connection between them and terms like "necessary connection," "power," and "force" do not mean what philosophers have typically thought. Rather than conclude that such terms are meaningless, Hume develops his own positive account, suggesting that they are based on a subjective feeling of the mind to infer the effect from a given cause, a feeling or expectation that arises only after repeated experience. Because our idea of causality is based on an impression of an internal custom or habit, it does not actually contain any necessity, but is rather restricted to contingently occurring constant conjunctions.

Second, Hume argues that the ultimate foundation of our knowledge of causality is not *a priori* reason, but rather custom, habit, or experience (where these terms indicate nothing more than repeated instances of certain kinds of impressions we have encountered in the past).[4] For if we are presented with an object that we have never seen before, we are incapable of determining, on the basis of reason alone, what effects this object will have. Instead, experience must provide us with such information. Hume then points out that since our knowledge of cause-effect relationships is based on various impressions from the past, we are not in a position to justify causal laws that would necessarily hold in the future. To do so would require establishing what Hume calls the principle of the uniformity of nature, which asserts that nature will continue to act in the future as it has in the past. However, since our empirical evidence is limited to the past and since no contradiction would seem to arise if nature were to change (i.e., reason has no justificatory force on the question), the principle of the uniformity of nature cannot be justified and no inference to the future is warranted. In a final twist, Hume

---

he could have brought Hume's critical points to Kant's attention in their conversations, and, to make matters even less clear, one cannot necessarily count on Hamann for straightforward objective accuracy in representing Hume's views, since Hamann had a distinctive agenda that departed considerably from Hume's. A helpful historical discussion of this issue is Manfred Kuehn's "Kant's Conception of Hume's Problem," *Journal of the History of Philosophy* 21 (1983): 175–93. In this paper, I abstract from the complex historical details in order to focus on a systematic comparison of Kant's and Hume's positions.

[2] Various medieval scholastics (e.g., Aquinas and Suarez), modern "rationalists" (e.g., Spinoza and Leibniz), and even the occasional "empiricist" (such as Locke in certain passages) had attempted to maintain powers, forces, or necessary connections.

[3] Hume presents these arguments in §§2 and 7 of the *Enquiry Concerning Human Understanding*.

[4] Hume presents this argument in §§4–5 of the *Enquiry*.

argues that, in spite of our lack of any rational or empirical justification, we still can and do expect the future to be like the past, given that it is simply part of human nature, whose fundamental basis is the passions, to make such inferences.

What has received by far the greatest amount of attention regarding Kant's response to Hume on the issue of causality is his Second Analogy of Experience, since it argues that causality, understood as involving some sort of necessary connection, is a necessary condition for knowledge of objective succession.[5] Accordingly, Kant's strategy would seem to be that if Hume were to admit something as minimal and basic as the fact that we have knowledge of objective succession (i.e., of one thing happening after another), then—granting the argument of the Second Analogy—Hume would thereby be implicitly committed to the very notion of causality and necessary connection that he had attacked so powerfully in the *Enquiry*.[6] Commentators on Kant's argument are divided, however, as to what its ultimate goal is.[7] Is Kant, as some think, simply attempting to justify the "every-event-some-cause" principle (thus responding to Hume's first skeptical argument by showing that every event requires a necessary connection or cause)? Or is he, as others suggest, undertaking the more ambitious task of establishing the "like-cause-like-effect" principle (therefore developing an answer to Hume's second argument by establishing the necessity of causal laws)?[8] Part of the reason for the lack of a consensus is that Kant slides back and forth from one principle to the other without even seeming to recognize any difference between them. But the problem is not merely exegetical; so far no one has presented a cogent philosophical refutation of Hume, one that he and his followers would be compelled to accept as defeating their position.

---

[5] Quotations from Kant's *Critique of Pure Reason* will be cited according to the standard A and B pagination for the first and second editions, respectively. Quotations from other works will cite the volume and page number of the Academy edition (*Gesammelte Schriften* ed. Königlich Preussische Akademie der Wissenschaften [Berlin: de Gruyter, 1902– ], volumes 1–29). Translations from the *Critique of Pure Reason* will be from Immanuel Kant, *Critique of Pure Reason*, Paul Guyer and Allen Wood, ed. & trans. (New York: Cambridge University Press, 1998). Translations from Kant's lectures on metaphysics will be from Immanuel Kant, *Lectures on Metaphysics*, Karl Ameriks and Steve Naragon, ed. & trans. (New York: Cambridge University Press, 1997). In a few instances, I have made minor modifications in these translations.

[6] This reconstruction of Kant's argument in the Second Analogy presupposes that Hume is *not* a complete skeptic, that is, for example, someone who is a skeptic about objective succession as well. Some commentators (e.g., P.F. Strawson) suggest that Kant has a further argument (the Transcendental Deduction, the Refutation of Idealism, or even the Second Analogy itself) to show that a more radical skeptic, who explicitly admits only the existence of his own mental states, is in fact implicitly committed to the distinction between the subjective and objective orders and thus to notions such as objective succession as well. Thus, the presupposition of the argument of the Second Analogy as reconstructed above may be discharged so that Kant would even be in a position to respond to a much more radical skeptical position.

[7] Cf. Lewis White Beck, *Essays on Kant and Hume* (New Haven: Yale University Press, 1978), Henry Allison, *Kant's Transcendental Idealism. An Interpretation and Defense* (New Haven: Yale University Press, 1983), Paul Guyer, *Kant and the Claims of Knowledge* (New York: Cambridge University Press, 1987), Michael Friedman, "Causal Laws and the Foundations of Natural Science," in Paul Guyer, ed. *The Cambridge Companion to Kant* (New York: Cambridge University Press, 1992), 161–99, and James Van Cleve, *Problems from Kant* (New York: Oxford University Press, 1999).

[8] Of course, hybrid interpretations are possible. For example, one common strategy is to establish the "every-event-some-cause" principle first and then use the idea that a causal rule implies generality in order to establish the "like-cause-like-effect" principle. See Lorne Falkenstein, "Hume's Answer to Kant," *Noûs* 32 (1998): 331–60.

Rather than engage in this well-known debate directly, I propose to consider how Kant is replying to Hume by focusing our attention on a slightly different question, one that has not been widely discussed, namely: What model of causality does Kant use in his reply to Hume? Typically, discussions of Kant's views on causality assume that, except for the disagreement on the question of necessary connections, Kant accepts Hume's model of causality, according to which one determinate event (e.g., the motion of one billiard ball at one moment in time) causes another determinate event (e.g., the motion of another billiard ball at a later moment in time). After all, if Kant were to employ a model that differed from Hume's in significant ways, how could he possibly hope to refute Hume's position? However, the justification for this natural presupposition about Kant's model is not exclusively philosophical. Since Kant himself asserts that it was Hume who first awoke him from his dogmatic slumber and, for that matter, precisely on the issue of causality, textual evidence, too, supports the idea that Kant accepts Hume's model of causality.[9]

In this paper I argue for the contrary hypothesis that Kant's model of causality is not that of one determinate event causing another determinate event, as Hume held, but rather involves both causal powers, which Kant understands in terms of unchanging grounds (since a causal power provides the reason or ground for its effects), and the exercise of these powers, which Kant describes as a temporally indeterminate activity that brings about (or posits) determinate states of substances that are therefore passive in this respect. After briefly discussing Hume's understanding of events and how they fit in with his model of causality and his more general philosophical project, I argue (1) that while Kant's Second Analogy only suggests that there may be important differences between Kant's and Hume's models, the Third Analogy of Experience is inconsistent with any Humean event-event model of causality, whether simple or complex.

I then argue (2) that Kant's model of causality ought to be understood positively in terms of grounds or causal powers by describing several fundamental features of grounds (as they contrast with events), by showing how a model based on grounds can avoid the problems faced by event-based models, and by pointing out that the crucial feature of grounds is the way in which they actively determine substances that are related to each other. At this point, I attempt (3) to describe in greater detail some of the unique features of Kant's model of causality by focusing on his discussion of the law of continuity. According to it, Kant's model of causality involves a cause bringing about a change from one temporally determinate boundary-state to another—a change he calls the effect—by means of a temporally indeterminate activity, whose function it is to bring about the determinacy of the effect. Passages from transcripts from Kant's metaphysics lectures explain further and justify the crucial features of this model, revealing in particular the irreducibility of the causal relation and the importance of its dichotomy between activity and passivity.

---

[9] For a fuller assessment of the historical evidence (which does not lend support to this claim, but suggests rather that on the issue of causality Kant was focused more on Leibnizian-Wolffian considerations), see my "The Development of Physical Influx in Early Eighteenth Century Germany: Gottsched, Knutzen, and Crusius," *Review of Metaphysics* 49 (1995): 295–339, and "Kant's Theory of Physical Influx," *Archiv für Geschichte der Philosophie* 77 (1995): 285–324.

To show how Kant's abstract model of causality can be made more concrete and the generic notion of activity employed in it illustrated in specific ways, I consider how he fills out his model in two prominent cases. First (4) I show that Kant's appeal to Newtonian attractive and repulsive forces in the *Metaphysical Foundations of Natural Science* nicely illustrates his abstract model of causality. Unsurprisingly, however, it turns out that appealing to Newtonian forces brings with it an empiricist implication, namely that such forces cannot ultimately be used to generate any independent meaning for the notion of activity that is a distinctive and central element of Kant's model of causality. Then (5) I argue that Kant's best strategy for providing an intelligible explanation of this notion of activity lies in an appeal to our consciousness of our own mental activities in self-consciousness, a mode of awareness that is both clear and as available to Hume as it is to Kant. Armed with this detailed understanding of Kant's model of causality, I am in a position to reconstruct (6) an argument, based on Kant's notion of a ground, that would both establish the necessity of causal laws and render intelligible why he could have switched back and forth between weaker and stronger causal principles without feeling the need to draw our attention to their differences.

Finally (7) I argue that while one might still be tempted to view Kant as replying to Hume directly or else to reinterpret Kant's model in such a way that one could find a refutation in it of Hume's skeptical doubts about causality, the fundamental differences between Hume's and Kant's models of causality are in fact simply a subset of more radical contrasts in their broader philosophical projects, making such reinterpretations impossible. These differences thus entail that Kant's reply to Hume consists not in a direct *refutation* of Hume's skeptical doubts on Hume's own terms, but rather in Kant rejecting Hume's position and advancing his own distinctive model of causality as a *competitor*. Viewed in this light, one can see more clearly that Kant's model of causality not only occupies a sophisticated middle position between more extreme views, but also is a prime illustration of his novel philosophical methodology, which stresses the need for, and the utility of, a particular kind of epistemological argument that can establish metaphysical doctrines he found attractive throughout his career. What's more, the emphasis he places on the notion of activity and determination in his model of causality is consistent with some of the distinctive features of his broader philosophical system, including that of autonomy, which he understands as a special kind of self-determination. As a result, pursuing a different understanding of Kant's model of causality both in abstract form and in several of its concrete instantiations displays promise for a reevaluation of the virtues that attach to Kant's philosophy as a whole.

## 1. HUME AND EVENT-EVENT MODELS OF CAUSALITY

Before we turn our primary focus to determining what Kant's model of causality is, consider briefly how the first of Hume's skeptical arguments discussed above relates to both his model of causality and his larger project. As we saw, one central assumption of Hume's argument for the distinctness of cause and effect is that we have no impression of a necessary connection between the event that we consider

the cause and the event that is thought to be its effect. What can Hume say in support of this assumption beyond providing a list of examples that he finds, in his own case, to be consistent with it?

Hume's argument for the distinctness of cause and effect could be primarily epistemological in character and rely simply on the fact that due to the coarseness of our senses we do not happen to have the ability to perceive necessary connections or "secret powers" in nature. We are built in such a way that we are able to perceive colors, motions, sizes, shapes, etc., but not necessary connections. Just as some species of animals are color-blind, we happen to be "power-blind." Accordingly, even if cause and effect were somehow necessarily connected in nature, we would still perceive them as distinct, since we do not possess the ability to have an impression of a necessary connection.

However, Hume can also endorse a stronger argument for the distinctness of a cause from its effect, one that is based on his ontology of events.[10] In his discussion of the infinite divisibility of space and time in Book One Part II of *A Treatise of Human Nature* Hume commits himself to the view that an event is a state of an object at a particular moment in time that does not endure or span any length of time.[11] Put informally, for Hume we experience the world by means of completely discrete, instantaneous mental snapshots. In filling out this view, he argues that we get our idea of duration from our idea of succession, and our idea of succession is in turn derived not from an impression of a change of state of an object, but rather from noticing or feeling the succession of distinct impressions in the mind.[12] In short, according to Hume we must get our idea of succession from a succession of impressions rather than from an impression of succession, since an impression of succession would necessarily endure over time. But if an event is an instantaneous state of an object at a particular moment in time and if a cause must precede its effect in time (as Hume stipulates), then the cause and the effect must exist at wholly different times and thus be completely distinct from each other.[13] According to this line of argument, causes and effects, as events that we perceive, are not merely contingently but rather necessarily distinct for Hume.

---

[10] Although this argument is broadly epistemological in nature, just as the argument from the coarseness of our senses is, there is a sense in which it reveals Hume's most basic ontological commitments in a way that the argument from coarseness does not. Also, the previous argument establishes only the contingent distinctness of cause and effect, whereas this argument establishes necessary distinctness, making it more powerful.

[11] See David Hume, *A Treatise of Human Nature*, L.A. Selby-Bigge, ed., P.H. Nidditch, rev. (Oxford: Clarendon Press, 1978), 26–33. Hume speaks of "quality" in this context, but, in light of his overall project, it is plausible to think that simple, indivisible qualities are to be understood as events. While Kant may not have been aware of this section of the *Treatise* when writing the *Critique of Pure Reason*, this conception of events can be naturally read off from Hume's treatment of these issues in the first *Enquiry*. See, for example, Hume's conclusion that "In a word, then, every effect is a distinct event from its cause," David Hume, *An Enquiry Concerning Human Understanding* (Indianapolis: Hackett Publishing Company, 1977), 19. Moreover, if Kant was not aware of Hume's actual position, then this fact ends up providing additional support for understanding Kant's reply to Hume as I do below.

[12] See Hume's *Treatise*, 34–37.

[13] §7 of the *Enquiry* directly supports such a claim. This argument also presupposes the idea that events that occur at different times cannot be related by means of a necessary connection. Hume may ultimately have to appeal to the second skeptical argument presented above to justify this presupposition.

Moreover, Hume's understanding of an event as an instantaneous state at a particular moment in time is not an accidental feature of his overall position. For his understanding of events as discrete, instantaneous entities is a fundamental assumption in his project of trying to build our complex view of the world out of the simplest and most basic building blocks to which we have immediate access. Since an event that endures is necessarily divisible, it could not form a *simplest* building block with which one could construct a secure edifice. Thus, in attempting to construct the causally connected everyday world we know out of such instantaneous events, Hume realizes quite clearly that there could be no objectively necessary connections that might serve as "the cement of the universe" and thus that he is limited to purely contingent relations between these building blocks (e.g., constant conjunctions) along with the subjective feelings or expectations to which they give rise. Hume's understanding of events, his model of causality, and his general project thus form a cohesive and powerful whole.

With a few of the main contours of Hume's position on causality in mind, we can now start to consider what Kant's model of causality must be like, beginning with the claim of the Second Analogy of Experience and the structure of his main argument for it.[14] The Second Analogy asserts that a cause (or perhaps a causal law) is required to bring about the succession of states in a substance as its effect. Kant's argument is based on the idea that since the subjective temporal order in which we happen to apprehend the states of an object in intuition does not directly reflect the temporal order of states in the object, only a causal connection could determine that one state of the object actually occurs after another state of that object, that is, that an event has occurred. Kant uses his famous example of the ship and the house (A192/B237–A193/B238) to illustrate both the distinction between the subjective and objective temporal orders and the fact that no inference from the former to the latter is valid unless it employs the category of causality.

Now two points in the Second Analogy are directly relevant to our current question. First, it makes clear that Kant understands an event (*ein Ereignis* or, less frequently, *eine Begebenheit*) not as a state of an object at a particular moment in time (as Hume does) but rather as a *change* of an initial to a later state of an object.[15] What's more, such a conception of events plays an essential role in the argument of the Second Analogy.[16] Not only does the argument try to show that causality is a necessary condition for our knowledge of a change of states in an object, that is, for our knowledge of an event, but it is also specifically a *change* of state that is required for the argument to work, since what causality is required for is precisely the *change* of the successive states of an object.[17] As a result, a Humean event does not possess the feature that is required by Kant's argument.

---

[14] What I say here should be consistent with any interpretation of the Second Analogy.

[15] Kant uses two separate terms '*Wechsel*' and '*Veränderung*' in such contexts, which are typically, though not always, translated as 'change' and 'alteration.' The idea is that states change while the object having the changing states alters in the process. Except in certain contexts (e.g., in commenting on quotations in which alteration is used), I shall use 'change,' since the difference between change and alteration is of no significance for current purposes.

[16] I am claiming here only that Kant's conception of the effect necessarily involves two successive states, not that Kant has no room in his ontology for determinate states of an object.

[17] As we shall see below, Kant's Third Analogy of Experience does not require change, even if it is involved in the succession of the simultaneous states of two substances.

Second, it is quite striking how little information the main *argument* of the Second Analogy provides about what a cause must be. It requires neither that a cause be an event, nor that it be prior to its effect, nor even that it have any specific empirical characteristics at all. In the Second Analogy, the central task of a cause is simply to determine the states of a substance as successive. Thus, taken in isolation, it rules out neither Cartesian occasionalism nor Leibnizian pre-established harmony, since it is possible that the cause of change is either God or internal to the substance whose state is changing. As a result, instead of confirming the expectation one might initially have had, namely that Kant's model of causality is the same as Hume's (aside from their disagreement about the existence of necessary connections), the Second Analogy indicates that Kant conceives of the effect as a change of state over time (rather than as a state at a particular moment in time) and the cause in terms of what it can accomplish, namely the determination of the successive states of a substance (which simply does not speak to whether or not it could be an event). Accordingly, the argument of the Second Analogy does point to some differences between Kant's and Hume's models of causality, but it does not suggest that these differences would be especially divisive or that the two positions could not, perhaps, be reconciled.

Kant's Third Analogy of Experience, which is typically passed over in silence in discussions of Kant's views on causality, turns out to be much more informative for understanding his model of causality. The Third Analogy asserts that substances must stand in mutual interaction in order for knowledge of the simultaneity of their states to be possible. It is clear, on the face of it, that Kant's notion of mutual interaction expresses a more robust type of causality than the relatively sparse notion of causality required by the Second Analogy. For mutual interaction is a two-way causal relation, where each causal relation holds between a substance and the states of a substance that is distinct from it. But how exactly is such a two-way causal relation to be understood? Specifically, what must a substance be like in order to be able to determine the states of another substance (in the way that the Second Analogy suggests)? Does such determination require that an event occur in the substance that is the cause, as a Humean event-event model would hold? That is, can the model of causality prescribed by the Analogies be understood in terms of events?

In order to begin to determine the basic features of Kant's model of causality, take a Humean event-event model of causality and consider how it might be used to construct the Third Analogy's notion of mutual interaction. Attempting to understand mutual interaction in terms of a *simple* event-event model of causality generates an explicit contradiction. If causes and effects are instantaneous events and mutual interaction is a two-way causal relation, then mutual interaction would be a two-way causal relation between two events. That is, one event would cause a second event, which would, in turn, cause the first event, which is obviously contradictory. If a cause must be prior to its effect (as Hume asserts) and mutual interaction is a reciprocal causation between instantaneous events, then a single event would have to be both prior and subsequent to another event, which is clearly a contradiction. Since this contradiction arises due to the impossibility of one event being both *prior* and *subsequent* to another event, one could avoid it by simply rejecting the temporal asymmetry of causality.

However, a second contradiction immediately takes its place, one that does not involve the temporal priority of the cause over the effect and is thus problematic for any simple event-event model of mutual interaction. If mutual interaction consists in two events causing each other, then the problem arises that if the first event causes the second event, then the second event cannot in turn cause the first event, as mutual interaction would claim, since the second event depends on the first one for its very existence and is thus unable to be that upon which the first one depends for its existence. In other words, it is a contradiction to claim that one event is both the cause and the effect of another at one and the same time, because causality entails that the existence of the effect depends on the existence of the cause and it is impossible for one event to depend on another at the same time that the second event depends on it.[18] It is important to notice, however, that this contradiction arises for *any* simple event-event model of mutual interaction, that is, whether events are understood as instantaneous states of affairs at particular moments in time, as Hume would have it, or as changes of state, as Kant believes. That is, Kant's Third Analogy is incompatible not only with the model of causality that Hume actually holds, but also with any simple event-event model of causality.

What generates both of the contradictions for the simple event-event model of causality is the fact that in mutual interaction one event is supposed to be both the cause and the effect of another. In order maintain mutual interaction while also avoiding the contradictions just encountered, one must develop a more *complex* model of causality by splitting up each of the events that had served as the causal relata into distinct entities so that one and the same event does not perform both functions at the same time. As a result, in order for mutual interaction to be a coherent possibility, it is necessary that what one might call "the causal aspect" of a substance not be identical to "the effect aspect" of that same substance, where the causal aspect of the substance is that part of the substance by means of which one of the two relations constituting mutual interaction brings about its effect, and the effect aspect of that substance is the effect brought about by the other of the two causal relations constituting mutual interaction.

If a distinction between the cause and effect aspects of a substance is drawn in this way, then one can construe Kant's model in terms of events as follows. Event $e_1$ that occurs in one substance, causes event $e_2$ in another substance, while this second substance, by means of event $e_3$ that occurs in it, causes event $e_4$ in the first substance. Because event $e_1$ brings about event $e_2$ and event $e_3$ brings about event $e_4$, no event is both the cause and the effect of another and the contradictions faced by simple event-event models do not arise. However, does this more complex event-event model of causality represent a coherent interpretation of mutual interaction? There are two different basic versions of such a model. First, if one accepts the idea that a cause must precede its effect, no contradiction arises and

[18] Technically, a contradiction arises only if one accepts the further principle that nothing could be the cause of itself. Kant explicitly accepts such a principle in his *Nova dilucidatio* (1:394). This objection presupposes that we are talking about simple, distinct events. As we shall see presently, if one were to allow for complex events (i.e., events that had "smaller" events as their components), then this particular contradiction could be avoided.

mutual interaction at least represents a metaphysical possibility. For example, this model might be understood such that we would have event $e_1$ at $t_1$ causing event $e_2$ at $t_2$, and then event $e_3$ at $t_3$, in turn, causing event $e_4$ at $t_4$. While such a version is apparently coherent on its own, it is unacceptable in the context of an interpretation of the Third Analogy because mutual interaction does not, in that case, enable knowledge of coexistence. We have knowledge of the first substance only at $t_1$ and $t_4$ and knowledge of the second substance only at $t_2$ and $t_3$, and thus no knowledge of the states of both substances at the same time.[19]

In order to ensure that knowledge of simultaneity is established, one might propose a second version of this model as follows. Event $e_1$ at $t_1$ causes event $e_2$ at $t_2$, while event $e_3$ at $t_1$ causes event $e_4$ at $t_2$, with events $e_1$ and $e_4$ occurring in a first substance and events $e_2$ and $e_3$ occurring in a second substance. This modified model is still distinct from the simple event-event models since it distinguishes causal aspects of substances (events $e_1$ and $e_3$) from their effect aspects (events $e_2$ and $e_4$), and no contradiction arises in virtue of any reciprocal existential dependency. Further, unlike the original version of this model, because events $e_1$ and $e_3$ are both at $t_1$ and events $e_2$ and $e_4$ are both at $t_2$, the causal ties guarantee that we would be able to know the simultaneity of the two substances.

There is, however, a fatal difficulty with this version of this model. For using temporal indices, such as $t_1$ and $t_2$, smuggles in coexistence illegitimately. To illustrate this difficulty more clearly, consider the same model using the terms 'before' and 'after' in place of $t_1$ and $t_2$. Such a replacement is warranted on Kant's account of causality, since Kant's concern never extends beyond establishing the minimal notion of temporal *order* that is involved in succession (and not its measurable lapse, cf. A203/B248). Now, on the version of mutual interaction just proposed, the first causal tie does *not* determine that event $e_1$ at $t_1$ causes event $e_2$ at $t_2$, but rather that event $e_1$ is *before* event $e_2$. Similarly, the second causal tie determines that event $e_3$ occurs before event $e_4$. But stated in this manner, coexistence has not been established between any of the events and the previous version's difficulty reappears. First, it has not been shown that events $e_2$ and $e_4$ coexist, but rather only that they occur after events $e_1$ and $e_3$, respectively. Second, it has not been shown that events $e_2$ and $e_4$ occur an equal temporal distance after events $e_1$ and $e_3$. It is entirely possible on this model that the one later event ($e_2$) occurs long after its causally related initial event ($e_1$), whereas the other later event ($e_4$) occurs just a split second after its causally related initial event ($e_3$) so that one cannot infer the simultaneity of the later events ($e_2$ and $e_4$) from their occurring at an equal temporal distance after the initial events ($e_1$ and $e_3$). Third, even if one could determine equal temporal distances between both sets of initial and later events, one could not infer the simultaneity of the later events from this fact, since this inference requires the simultaneity of the initial events, which has also not been shown. Therefore, one initial event, which is determined to be prior to

[19] C.D. Broad, *Kant: An Introduction* (New York: Cambridge University Press, 1978), appears to hold a version of such an interpretation. Consider his statement: "to say that two substances A and B are in mutual interaction would seem to have the following meaning. Every alteration $a_1$ in A causally necessitates a *later* alteration in $b_1$ in B; this in turn causally necessitates a *later* alteration $a_2$ in A; this causally necessitates a *later* alteration in B; and so on" (178, emphasis added).

one later event, is not necessarily simultaneous with the other initial event, which is determined to be prior to the other later event.

Faced with another dead end, one might think that the source of the problem lies in introducing temporal asymmetry into the causal relations. Accordingly, if one rejects the idea that a cause must be prior to its effect (as Kant clearly does at A203/B248), then one might think that the temporal disparity between the events could be avoided in such a way that mutual interaction could still be necessary for the simultaneity of the two substances without entailing any contradiction. The most promising complex event-event model now is as follows. Event $e_1$, in the first substance, at $t_1$ causes event $e_2$, in the second substance, at $t_1$, which in turn causes event $e_3$, in the first substance, also at $t_1$. Since i) events $e_1$, $e_2$, and $e_3$ all obtain at $t_1$, ii) the simultaneity of the two substances can be known, and iii) there are two causal relations going in opposite directions, it would be natural to describe them as an instance of mutual interaction that also allows for knowledge of the coexistence of substances.

However, this model faces two new objections. First, if this model were possible, it would not so much support the Third Analogy, as rather be a devastating objection to it. For if it were possible to assert that event $e_1$ at $t_1$ causes event $e_2$ at $t_1$, then there would be no need to assert that event $e_2$ causes event $e_3$ at $t_1$ in order to establish the simultaneity of the two substances. That is, there would be no need to assert *mutual* interaction between the two substances, since we would already know, on the basis of the first causal relation, that the two substances coexist at $t_1$.[20] Second, and even more seriously, this model presupposes a feature of the causal relations making up mutual interaction that is prohibited by the argument of the Third Analogy. In particular, it presupposes that a substance can determine its own place in time by assuming that event $e_1$ occurs in the first substance at $t_1$. As is clear from the argument of the Third Analogy, a substance cannot determine its own place in time and therefore requires the causal efficacy of a substance distinct from it, which ultimately generates the need for mutual interaction (since each substance requires its place in time to be determined by another).[21] Because the first substance cannot determine its own place in time, it cannot determine that event $e_1$, which causes event $e_2$ at $t_1$, occurs at $t_1$. Yet nothing else could determine event $e_1$ at $t_1$ either since the defining feature of the complex event-event model was its separation of cause aspects from effect aspects. Thus, complex event-event models of causality fail as an interpretation of Kant's model of causality, just as simple event-event models did.

---

[20] Paul Guyer (*Kant and the Claims of Knowledge*, 272–73) criticizes Kant's argument in the Third Analogy as being unable to establish *mutual* interaction, since a simple causal relation of the sort described above would suffice for knowledge of coexistence. While Guyer and I agree about the force of this criticism, we disagree about what follows from it. He infers that the argument of the Third Analogy fails, whereas I take the objection as motivation to look for a philosophically defensible way of understanding mutual interaction such that this objection does not arise.

[21] See my "Kant's Third Analogy of Experience," in *Kant-Studien* 88 (1997): 406–41, for a detailed reconstruction of Kant's argument in the Third Analogy.

## 2. GROUNDS AND CAUSAL POWERS

If Kant's model of causality cannot be explained solely in terms of events, what other options are open to him? In particular, since Kant explicitly identifies the effect with an event (as a change of state), the decisive question must be: What does Kant think that a cause is? The First Analogy of Experience might seem to provide an immediate and simple answer to this question, since it contains an argument for phenomenal substances, and it might seem obvious that a cause must be a substance.—What causes the motion of the second billiard ball? The first billiard ball, which is simply a spatial substance.—While there is a non-trivial sense in which such an answer is correct, it is important to recognize that it can be only a small part of Kant's full answer. For this position is subject to a series of pressing philosophical questions, questions that one cannot answer by appealing simply to the mere notion of a substance as such: How can the *mere existence* of a thing bring about an effect? How is one to understand that such a cause would bring about its effect *at any one time* rather than at any other, if it is supposed to be the thing rather than its state at any given time that is the cause? How does the mere existence of *one* thing explain an effect in *another*? Is it not the *state* of the thing at a particular time (as opposed to the *thing* per se) that could explain the effect?[22]

But note that the argument of the previous section has ruled out that the cause could be what these questions might suggest, namely a determinate state of the substance. In terms of our concrete example, if these questions show that the cause of the motion of the second ball cannot simply be the first billiard ball per se, Kant's argument in the Third Analogy shows that it cannot be the motion of the first billiard ball either. But if the cause cannot be simply a substance (the first billiard ball) nor a determinate event in it (its motion), one faces, once again, the question of what it can be. It is helpful to note here that Kant sometimes uses the phrase "the causality of the cause" and, on several occasions, explicitly distinguishes between the cause and "the causality of the cause."[23] This suggests that it is Kant's notion of the causality of the cause that is crucial to understanding his model of causality. But how is "the causality of the cause" to be understood? To answer this question, let us first attend to Kant's notion of a ground and then consider how his model of causality can be explained in terms of such a notion.

The basic idea of a model of causality for which grounds are central is that one substance determines the successive states of another by means of an unchanging

---

[22] These questions are simply reformulated versions of the very questions that proponents of event causation (such as Davidson and Broad) pose to advocates of agent causation (such as Reid, Taylor and, for most of his career, Chisholm). It is clear that Kant is aware of such questions as early as his *Nova Dilucidatio* in 1755, where he explicitly distinguishes between the existence of a substance and its causal relations with other substances.

[23] See, for example, his Metaphysics Mrongovius lectures: "Causality is the determination of a cause by which it becomes a cause, or the determination of the relation of a thing as cause to a determined effect. Thus the cause is always to be distinguished from the causality" (29:893) and the $L_2$, lectures, where he notes: "When the cause has been posited, the effect is posited [*posita causa ponitur effectus*] already flows from the above. But when the cause has been cancelled, the effect is cancelled [*sublata causa tollitur effectus*] is just as certain; when the effect has been cancelled, the cause is cancelled [*sublato effectu tollitur causa*] is not certain, but rather the causality of the cause is cancelled [*tollitur causalitas causae*]" (28:573).

ground that constitutes its essential nature. Since a ground is both an essential feature of a substance and a source of change (insofar as it determines the successive states that constitute change), it cannot itself change from one determinate state to another (as that would entail an infinite regress). As a result, a ground is not temporally determinate in the way in which the effect is (since the effect, unlike the cause, has one determinate state at one moment in time at its beginning and another such state at its end).

One can find three distinct lines of support for understanding Kant's model of causality in terms of grounds. First, understanding Kant's model of causality in terms of grounds allows one to avoid the problems that the various event-based models faced in explaining how mutual interaction is to be possible. Second, the structure of Kant's explanation of motion in terms of grounds in his pre-Critical period is analogous in fundamental respects to what is needed to account for the knowledge of simultaneity discussed in the Third Analogy. Finally, upon closer inspection, one can find unambiguous textual evidence in the Second Analogy that Kant adopts precisely this notion of ground. In addition, if one acknowledges that grounds are nothing other than causal powers, one can see further textual support for the claim that Kant's model of causality involves grounds, and also understand why Kant would not have thought it necessary to emphasize his divergence from (Humean) event-event models.

Consider how grounds differ from Humean events and how understanding Kant's model of causality in terms of them can avoid the problems that the various event-based models faced. Kant's grounds are distinct from Humean events in several respects. First, unchanging grounds endure throughout the change of states that they cause, whereas instantaneous events pop into and out of existence. Second, since grounds determine changing states, they are not ontologically distinct from their effects in the way that events are from each other. Finally, whereas events are temporally determinate (since they occur at a determinate instant in time), grounds are temporally indeterminate, given that they do not change from one determinate state to another.

Now recall the various difficulties that event-based models encountered in attempting to account for simultaneity by means of mutual interaction. First, for simple event-event models, the cause and the effect depended on each other ontologically. Second, for complex event-based models (which distinguish between cause and effect aspects of a substance), the cause and effect aspects of the one substance could diverge temporally from, and thus not be simultaneous with, the cause and effect aspects of the other substance. The cause of this potential lack of simultaneity was that the two causal relations that were to constitute mutual interaction could obtain independently of each other. The independence of these two causal relations was implied, in turn, by the fact that the cause-effect relationship used to construct mutual interaction invoked nothing more than events, and events, as Hume argued, are necessarily distinct from each other. Third, because of this second difficulty, it appeared that event-based models of causality could be used to explain simultaneity only if a substance could determine its own place in time (given that the second difficulty eliminated the possibility that other substances could do so), but this principle contradicts a fundamental assumption of the ar-

gument of the Third Analogy (since if a substance could determine its own place in time, then *mutual* interaction would not be necessary).

If Kant's model of causality is based on grounds rather than events, it can avoid the problems encountered by the different versions of event-event models as follows. First, since the ground of one substance that determines the (successive or changing) states of another substance does not, in turn, depend on those successive states, this model is obviously not committed to reciprocal existential dependencies of the sort that plagued simple event-event models. Second, if the way that the grounds of one substance determine the successive states of another is not independent of the way that the grounds of the second substance determine the successive states of the first substance, then it is possible that these grounds jointly determine their respective states. Since simultaneity is simply a particular instance of the joint determination of the states of substances, a model of causality based on grounds is not immediately barred from providing an explanation of simultaneity. Finally, since grounds are temporally indeterminate, there is no need to assume the temporal determinacy of the cause in the first place and one is thus not in danger of violating any fundamental assumptions of Kant's argument in the Third Analogy.[24]

Understanding Kant's model of causality in terms of grounds can also allow one to notice and then make use of parallels it has to his pre-Critical account of causality. In his *Nova dilucidatio* (1755), Kant argued for the principle that mutual interaction (which is defined in terms of what he calls "determining grounds") is required for mutual changes of substances, and one can use the case of motion to illustrate this principle. Suppose one body is changing its motion with respect to another. According to Kant's understanding of the principle of sufficient reason at this time, there must be a ground that causes this change (or the determinations that constitute it). Yet Kant argues that because grounds are unchanging, the grounds in the first body cannot cause a change in its own determinations. As a result, a ground in the second body must be the cause of the change in the first body. However, since motion is a reciprocal relational property, (a change in) the motion of the first body toward the second necessarily implies (a change in) the reciprocal motion of the second body toward the first. By reason of parity, if the grounds in the first body cannot cause its own (change of) motion, then the grounds in the second body cannot be the cause of its own (change of) motion either. As a result, a ground in the first body must be the cause of (the change of) motion of the second body, just as a ground in the second body must be the cause of (the change of) motion of the first body. That is, the mutual change of state of two substances is possible only if they stand in mutual interaction (i.e., if each one grounds the motion of the other).

However, describing the case of motion in this way is somewhat misleading, since it suggests that there are two separate events, the motion of the first body and the motion of the second, which require two independent causes, the ground in the second body and the ground in the first body. What must be emphasized

[24] Nor is one faced with the infinite regress that arises if were one to attempt to account for the temporal determinacy of grounds by appealing to further grounds.

about this example is that neither the motions of the two bodies nor the way in which their grounds bring about these motions are independent of each other. First, the motion of the one body toward the second both logically implies and is logically implied by the motion of the second body toward the first. Second, since a substance cannot contain the grounds of its own changes, one must attribute the ground of the change that each substance undergoes to the other substance. Yet, due to the reciprocal relations between the effects, there must be reciprocal relations between each of the grounds so that they are not independent of each other. Thus, what makes mutual interaction mutual is not merely the fact that both substances must make a causal contribution, but also the fact that they must do so together. In short, two grounds must *jointly* determine the reciprocal motion of the bodies with respect to each other.

To understand how grounds can jointly determine the states of substances, it is important to distinguish between grounds *simpliciter* and the way in which they *actively* determine the states of a substance *in a specific situation*. Grounds *simpliciter* are essential features of the substances they constitute, and are therefore as independent from each other as self-subsistent substances are. However, they are not essential features *insofar* as they *actively determine* the states of another substance. As early as 1762 Kant understood that (causal) relations between substances cannot be logically necessary, since that is incompatible with the independence that defines substances as such. Accordingly, there must be a distinction between grounds that are necessary for a substance to exist at all and the particular way in which they bring about effects or determinations in other substances. In order for grounds to determine the states of other substances (as opposed to simply constituting the subsistence of their own substances), they must i) be placed in some relation to those substances so that they can then ii) actively bring about specific determinations in those substances.[25] To revert to the case of bodies, whether two bodies are placed two or three feet away from each other makes a difference not in the essential grounds of the substances so placed (e.g., in their masses), but rather in the determinations that they bring about in each other (e.g., in how strongly they attract each other). But given that one and the same substance could be placed in different relations to different substances, there are different ways in which its grounds can be effective. That is, depending on how two substances are situated with respect to each other, their grounds can determine different states in them.

This description of how Kant's pre-Critical account of causality in the *Nova dilucidatio* invokes mutual interaction to account for motion allows us to see how Kant understands mutual interaction in his Critical period such that it is necessary for knowledge of simultaneity. First, it is clear that simultaneity is a reciprocal relational property just as motion is. If the state of one substance is simultaneous with the state of a second substance, then it is impossible that that state of the second substance not be simultaneous with the state of the first. Second, if changes in a substance require the causal efficacy of a distinct substance on the grounds

[25] Kant's way of putting this would be to say that substances must stand in some relation to each other if they are to belong to one and the same world.

that a substance cannot act on itself so as to change itself, then it is analogous to claim that since a substance cannot determine its own place in time, one must attribute the source of such determination to another substance. Third, just as one must appeal to both the masses of bodies and their relations (e.g., distance) to determine what motions they will cause, so too one must consider how both the grounds and the relations between grounds determine whether two states are simultaneous or not. Fourth, since motion, as a reciprocal relational property, must be determined by two separate but jointly determining grounds in both bodies due to the fact that no single substance can determine motion by itself, it follows, by reason of parity, that two grounds must jointly determine the simultaneous states of these substances as well.[26] Finally, if one must draw a distinction between grounds *simpliciter* and grounds insofar as they actively determine the states of the substances to which they are related, it becomes more understandable what Kant means when he distinguishes between a substance per se and that feature of a substance by means of which it can be a cause, that is, between the mere existence of a substance and the "causality of the cause" (i.e., the causality of the substance that is the cause). As a result, the analogies between Kant's explanation of the motion of bodies in his pre-Critical period and the basic elements of his Critical model of causality help us to see, at least in rough outline, what these fundamental features must be like.

---

[26] One might, however, still be concerned that this new model of mutual interaction is not ultimately able to avoid the problems faced by the event-event models. For if it is the case that Kant's notion of mutual interaction requires that the grounds of two substances jointly determine their simultaneous states, it might seem to follow that these grounds must be simultaneous just as their states are. If that is right, then does it not follow that they would also be temporally determinate and does that not entail, in turn, that this model too, despite the best of intentions, would encounter the very same difficulties that the simple and complex event-event models did? Moreover, this question might seem all the more pressing, since it could be strengthened by drawing on one crucial line of argument from the *Nova dilucidatio*. There Kant had attempted to show that i) in light of the principle of determining ground, changing determinations require changing grounds and ii) since grounds cannot change, change within an isolated substance is impossible. Given this line of argument, one might think that just as changing determinations require changing grounds, so, too, simultaneous determinations would require simultaneous grounds, and if grounds are simultaneous, they must also be temporally determinate. But in the Critical period Kant rejects the inference from the determinacy of (simultaneous) states to the determinacy of the grounds that posit them. If the grounds of simultaneous states were simultaneous and thus temporally determinate, then, given that the problem of time-determination would apply to the grounds of states as much as it applies to the states themselves, something would have to determine them. But if grounds were to require further grounds in order to be determinate, an infinite regress would ensue, since it follows that the former could be determinate only if the latter were, and they, in turn, could be determinate only if there were yet a further set of grounds, etc. If the inference from determinate states to determinate grounds were supported by a *metaphysical* principle, the problem would not arise in this form, since the determinate states could function as the grounds of the determinacy of the grounds. But it is clear that such a metaphysical principle is unacceptable in this context, since that would mean that the states would have to be the grounds of their very own grounds, which is obviously circular, and one would, once again, have reciprocal existential dependencies. It is much more plausible to read the inference from determinate states to determinate grounds as based on an epistemological principle (according to which the determinacy of the states merely indicates the determinacy of the grounds), but then the problem of time-determination, which is, at least in part, a metaphysical problem, has not been resolved. Thus the inference from the determinacy of simultaneous states to that of simultaneous grounds ought to be rejected.

Finally, lest one think that the Critical Kant entirely rejects (as dogmatic) his pre-Critical notion of an unchanging ground, note first that Kant explicitly endorses several of the central features of this notion in a relatively neglected passage from the Second Analogy: "According to the principle of causality *actions* are always the primary *ground* of all change of appearances, and *therefore cannot lie in a subject that itself changes*, since otherwise further actions and another subject, which determines this change, would be required" (A205/B250, emphases added).[27] That is, Kant argues specifically that any change requires a ground, and that the ground of any change cannot itself change on pain of infinite regress.[28] Moreover, it is neither the mere existence of a substance nor any determinate event in it, but rather *action* or *activity* that serves as the *ground* of the changes that objects undergo.

However, this is not the only textual evidence in the Second Analogy for attributing to Kant a model of causality that is based on this notion of a ground, though the other pieces of evidence are a bit more indirect, since they presuppose that the notion of a ground that Kant appeals to—that of an unchanging ground that brings about determinate change in other objects—is to be explained ultimately in terms of the notion of a causal power or force (*Kraft*), as understood in its most *basic* form by, e.g., Aristotle, Leibniz, and Locke. While it is true that Kant does not highlight the fact that his model of causality should be understood in terms of grounds or causal powers in the first *Critique*, two of his remarks in the Second Analogy explain why he would have thought there to be no need to do so. First, "where there is action, consequently activity and force [*Kraft*], there is also substance" (A204/B250), where '*Kraft*' could just as easily have been translated as 'causal power.' Second,

> This causality leads to the concept of action, this to the concept of force [*Kraft*], and thereby to the concept of substance. Since I will not crowd my critical project . . . with analyses that address merely the elucidation (not amplification) of concepts, I leave the detailed discussion of these concepts to a future system of pure reason—especially since one can already find such an analysis in rich measure even in the familiar textbooks of this sort. (A204/B249)

As it turns out, transcripts from Kant's metaphysics lectures provide a rich set of detailed comments on such "familiar" analyses, which reveal quite distinctive features of his understanding of causal powers. However, the immediate point to note is that Kant *takes for granted* that we can appeal to the notion of a *Kraft*, that is, causal power, and that his intended audience would already be sufficiently familiar with the broadest outlines of this kind of model of causality from standard metaphysics textbooks. Since virtually all philosophers in the modern period (except, perhaps, Hume) accepted causal powers, there would be no special need to

---

[27] Since the argument against event-based models of causality was based on explaining simultaneity and thus on the Third Analogy, one might think that Kant's model of causality in the Second Analogy is still similar to Hume's (even if the Third Analogy's is not). However, this passage (along with others, e.g., at B111, A194/B239, and A199/B244) from the Second Analogy reveals that Kant's commitment to the same model of causality in the Second Analogy as well.

[28] Kant's position in the Antinomies reveals that he does not object to the very idea of an infinite regress of determinate states. Rather, he objects to the idea that the *ground* of a change could itself change.

provide a basic analysis of the notion, especially since Kant has more pressing issues on his agenda in the *Critique of Pure Reason*.

## 3. THE PRINCIPLE OF CONTINUITY, ASYMMETRY, AND ACTIVITY

If the evidence described above suffices to establish that Kant's model of causality is to be understood in terms of grounds (and causal powers), we can now consider it in greater detail and attempt to state some of its most distinctive features more clearly by investigating a series of lesser-known, but highly significant passages. First, it can be helpful to see a more detailed description of how a cause is supposed to bring about its effect. Kant's portrayal of the principle of continuity is particularly helpful in this regard. Second, several passages from Kant's metaphysics transcripts reveal that his model employs an asymmetrical and irreducible relation between cause and effect. Significant argument is required in order to appreciate why the relation must be irreducible. Finally, these same texts also disclose that the asymmetry and irreducibility of causal relations is to be understood in terms of an active-passive distinction, a distinction that requires careful elucidation.

Consider first Kant's discussion of the principle of continuity. He sometimes refers to the principle or law of continuity as the principle of no leap (A228–29/B281, 28:41, 28:199–205, 28:662–63, 29:862, and 29:1006), and he repeatedly and explicitly endorses it throughout his Critical period: in the first *Critique*, in the *Metaphysical Foundations of Natural Science*, and in various transcripts from his metaphysics lectures.[29] Though Kant distinguishes different versions of the principle of continuity (especially in the metaphysics transcripts), the version that is relevant for our purposes concerns alteration, and states simply that all alteration is continuous. Kant's preferred way of illustrating the principle is with the example of the motion of a body: "when one body transfers from one point to another, then it must go through infinitely many intermediate spaces, it must go through all intermediate locations lying between the one point in the line and the other" (28:201). However, in several passages (e.g., 29:864), Kant makes it clear that the principle is relevant to any kind of alteration, not just the motion of a body.

In the Second Analogy Kant explains this version of the principle of continuity as follows:

> The question therefore arises, how a thing passes from one state = *a* into another one = *b*. Between two instants there is always a time, and between two states in those instants there is always a difference that has a magnitude. . . . Thus every transition from one state into another happens in a time that is contained between two instants, of which the former determines the state from which the thing proceeds and the latter the state at which it arrives. Both are therefore boundaries of the time of an alteration, consequently of the intermediate state between two states, and as such they belong to the whole alteration. Now every alteration has a cause, which manifests its causality in the entire time during which the alteration proceeds. Thus, this cause does not produce its alteration suddenly (all at once or in an instant), but rather in a time, so that as the time increases from the

[29] See also A208–9/B253–54 and 4:552–53.

initial instant *a* to its completion in *b*, the magnitude of the reality (*b*–*a*) is also generated through all the smaller degrees that are contained between the first and the last. All alteration is therefore possible only through a continuous action of causality, which, insofar as it is uniform, is called a moment. The alteration does not consist of these moments, but it is generated through them as their effect. (A208–9/B253–54)

This rich passage brings out a number of important points about Kant's model of causality. First, Kant thinks of an alteration as a transition from one determinate boundary state to another, with an infinite number of states in between the boundary states. Second, the transition from one determinate boundary state to another occurs not in an instant, but over time. More specifically, the alteration from one boundary state to another does not occur in discrete "jumps" but rather in a continuous flow. Third, Kant describes the effect as being brought about by means of a "continuous action of causality." In other words, the continuous change of states is brought about by a continuous activity of the cause. Finally, alteration does not consist of the continuous action of causality, but is rather generated through such action.[30]

These points suggest that we specify Kant's model of causality more precisely as follows. When a cause brings about its effect, it acts uniformly, thereby generating a continuous flow of states in an object from one determinate boundary state to another. While the states that constitute the effect are thus fully determinate, the causality of the cause, despite its activity, is uniform and does not change from one determinate state to another. Moreover, since temporal determinacy comes about only as the result of the causal activity, the "causality of the cause" or "the continuous action of causality" is not itself determinate and could never become determinate through the causal activity of a distinct cause.[31] For the only kind of entity that can be determinate is a *state* of an object that lies between other determinate boundaries; by contrast, the causality of the cause is not and cannot be a determinate state, but rather must be a continuously efficacious *activity*. The ontology Kant invokes in the context of filling out his model of causality is thus very different from Hume's ontology of distinct, instantaneous events.

---

[30] This passage confirms that Kant does not think of the causality of the cause as being noumenal. For this passage asserts that the causality of the cause can be continuous and uniform, which clearly implies some kind of temporality. Moreover, because Kant accepts (at A203/B248) that a cause must exist in order to bring about its effect and that the effect occupies a determinate span of time, the causal activity must continue to exist even after it has started to bring about its effect so that it can bring about the remaining part of its effect. If it is appropriate to say (as Kant does) that the causality of the cause *continues*, then it must be temporal and thus cannot be noumenal.

[31] It is especially important to be clear about this point. One can of course determine some state of the substance that was the cause, but this determination in no way implies that one has thereby shown that *that state* was the cause of the effect in question. Now, in the context of the Third Antinomy, Kant does say that "the causality of the cause through which something happens is always something **that has happened**" (A445/B482) and that "the causality of the cause of what happens or arises has also **arisen**" (A532/B560), statements one might naturally interpret as asserting that the causality of the cause is a temporally determinate state of a substance. However, in these passages (and similar ones in the Metaphysics Mrongovius transcripts, such as 29:923), all Kant needs to be committed to is the idea that "the causality of appearances rests on temporal conditions" (A532/B560). That is, when trying to explain how the conflict between free will and determinism arises for a transcendental realist, Kant needs to establish merely that the causality of a phenomenal cause *relies on* the existence of a temporally determinate, earlier state as a condition, not that the activity by which the effect is brought about *is* itself temporally determinate, since it is the existence of the earlier state that generates the deterministic regress of causes back in time and not the nature of the causal activity itself.

Kant's distinctive understanding of causal powers is also articulated in transcripts from his lectures on metaphysics from 1782–83, called the Metaphysics Mrongovius lectures. While these lecture transcripts often contain mere explications of Baumgarten's position, it is not unusual for them also to contain critical remarks that Kant is directing at Baumgarten's views from the perspective of his own Critical position. In the present context, they represent further clear textual evidence that Kant is committed to a causal powers model of causality, which thereby complements the philosophical and textual argument presented in the previous sections against event-event models of causality. More importantly, however, these transcripts contain materials that lay out Kant's distinctive understanding of causal powers insofar as they illustrate that for Kant i) whenever they are exercised, causal powers form an irreducible and asymmetrical relationship between the substances they reside in and the determinations or accidents that they produce and ii) the asymmetrical aspect of this relationship must be understood in terms of an active-passive dichotomy.

Kant's first task concerning causality in the Metaphysics Mrongovius lecture transcripts is to defend the idea that causality is based on a *real* or ontological ground-consequence relationship as opposed to Baumgarten's purely *logical* ground-consequence relationship. Kant's generic definition of grounds and consequences is uncontroversial, as it asserts that they are entities conjoined by the relation of connection (rather than that of opposition). If the one is posited, then the other is necessarily posited as well, and vice versa. If the connection is analytic (i.e., according to the principle of identity), then the ground-consequence relationship is logical and *a priori*, whereas if it is synthetic (i.e., if reason cannot comprehend the connection), then it is real and *a posteriori*.[32] Because the generic definition of ground and consequence entails that they are necessarily coexistent and completely symmetrical, Kant recognizes that one "cannot distinguish ground and consequence by [this generic] definition" (29:808). In order to capture the crucial asymmetry that holds between ground and consequence (namely that the consequence is posited *only because of* the ground and not vice versa), he refines his definition by means of a notion of determinacy as follows:

> The ground is that by which, having been posited, another thing is posited determinately, the consequence is that which is not posited unless something else is posited . . . for if there is a consequence, there must likewise always be a ground, and if something is a ground, there must likewise always be a consequence, but in the first case it is indeterminate, in the other determinate. (29:808)

Accordingly, the ground determines the consequence, but the consequence does not determine the ground in precisely the same way, since, in principle, a given consequence could follow from any one of a number of different grounds. Yet since the cause-effect relationship is described as an instance of the real ground-consequence relationship, which is asymmetrical precisely in virtue of the consequence being determinate and the ground being indeterminate, the effect must likewise be determinate and the cause indeterminate. This view is precisely what was required of Kant's model of causality by his argument in the Analogies.

[32] Surprisingly, Kant does not discuss in this particular context the possibility of a real connection that is synthetic and also *a priori*.

Does Kant's discussion of this asymmetry between the indeterminacy of the cause and the determinacy of the effect add to his notion of a causal power and the model of causality explicated in terms of it so far? Later in the ontology section of the Metaphysics Mrongovius lectures, Kant discusses the relational categories of substance-accidence (subject), cause-effect (principle), and active-passive (interaction). In describing the way in which accidents inhere in substances, Kant describes accidents as positive determinations of substances. His primary concern is to show that negative and logical predicates are not accidents in the same sense in which positive predicates are and that a substance does not carry its accidents in the same way that a bookcase supports its books. Yet he also clarifies the relationship between substance, accident, and determination by claiming that "insofar as a thing is determined positively, accidents inhere in it" (29:770). In short, accidents are simply positive determinations. Moreover, a substance cannot itself be a determination. For Kant holds that a determination must be a determination *of* something and a substance cannot be a determination of something if it is to be defined as "that which exists without being the determination of another" (29:770). So far, the picture is clear enough. The effect is determinate insofar as the cause brings about a change of determinate states, or accidents, in a substance. The cause, insofar as it is associated with a substance, must be indeterminate, because a substance, by definition, cannot be the determination of another.

But how do causal powers fit into this ontological framework? Kant continues in the Mrongovius lectures by relating his discussion of substances, accidents, and determinations to causality as follows: "With a substance we can have two relations: in relation to accidents it has power insofar as it is the ground of their inherence; and in relation to the first subject without any accidents, that is the substantial. Power is thus not a new accident, but rather the accidents are effects produced by the power" (29:770). Here Kant clearly asserts that a power is not an accident or determination; rather it must be some kind of relation between a substance and accidents, from which it follows that it is not itself determinate.

If Kant has thus stated clearly that a power is not an accident, then the temptation is to identify it with substance, as Baumgarten does. Again, Kant's discussion of Baumgarten's position in the Metaphysics Mrongovius lectures is instructive. While Kant agrees with Baumgarten that a (causal) power (*Kraft*) is "that which contains the ground of the inherence of the accidents" (29:771), he rejects Baumgarten's claim that power is identical to substance. "Since accidents inhere in each substance, he [Baumgarten] concludes that every substance is power. That is contrary to all rules of usage: I do not say that substance is a power, but rather that it has power, power is the relation of the substance to the accidents, insofar as it contains the ground of their actuality" (29:771). In this passage Kant clearly denies that a power is itself a substance. But, as we saw above, it could not be an accident either and it must thus be an irreducible relation between substance and accident. Moreover, Kant is clearly aware of this implication and explicitly embraces it. For he continues as follows: "We thus have something [power] that is not substance, yet also not accident" (29:771). Accordingly, Kant clearly rejects the assumption that everything is either a substance or an accident and asserts that a power is a relation between its substance and the determinate states it produces.

Does Kant have any motivation that extends beyond linguistic usage for understanding causal powers in this way? The first point to note is that one is committed to ontological entities that are neither substances nor accidents as soon as one accepts into one's ontology substances, accidents, and an inherence relationship between the two as more than conceptually distinct aspects of reality.[33] This commitment can be illustrated by considering how one might attempt to reduce the inherence relationship to accidents and substances. Is the inherence relationship an accident of the substance or a substance? If it is an accident, then it obtains only in virtue of the fact that it itself inheres in the substance. But of course this second inherence relationship would be an accident too and would obtain only if it stood in a third inherence relationship with the substance, ad infinitum. At the same time, it is equally problematic to claim that the inherence relationship is itself a substance, since if it were, then one would need an inherence relationship between it and the accident that is supposed to inhere in the first substance, but this inherence relationship, which would be a substance as well, would require yet another inherence relationship, etc. In this case, too, an infinite regress seems unavoidable. It seems that the only way to avoid an infinite regress lies in accepting inherence as an indeterminate relation between substances and their accidents.[34] That Kant understands this point is evident from the fact that he explicitly includes the substance-accident-inherence relationship as a *primitive relational* concept of the understanding in his table of categories.

But notice that the substance-accident inherence relationship is structurally analogous to the relationship between a cause and its effect as expressed by "the causality of the cause" or by the way in which a causal power grounds new determinations in another substance. For according to Kant, both substances and causes ground their consequences, namely the inherence of accidents. The primary difference for Kant is that a substance is an *inner* sufficient ground of its own accidents, whereas a cause is an *outer* sufficient ground of the accidents that are its effect.[35] That is, a substance is an inner ground for the inherence of its own accidents, while a substance is a cause insofar as it is an outer ground, i.e., a ground for the inherence of accidents in another substance. Accordingly, both substances and causes are real grounds and the main difference concerns whether the consequence, namely the inherence of accidents in a substance, is internal or external to the substance that contains the ground. Yet this difference does not immediately affect whether grounding is itself determinate or indeterminate. If grounding is an indeterminate relation between a substance and its accidents in the one case, it will be such a relation between a substance and its accidents in the other as

[33] One could avoid this line of argument by understanding accidents as simply ways that the substance has of existing. See A186–87/B229–30 for Kant's discussion of how accidents "relate" to substances. At the end of his discussion, he remarks that "this category also stands under the title of relations, but *more* as their condition than as itself containing a relation" (emphasis added).

[34] The same point holds true of objects, properties, and the exemplification relationship that holds between objects and their properties. Is exemplification a relational property? If so, it would entail an infinite regress of exemplification relationships, which would always be just further properties of which it would be unclear whether or how they related to the object itself. At the same time, it is difficult to see how a thing could literally be an exemplification.

[35] See 28:51–53 for Kant's most explicit discussion of this point in the Herder transcripts.

well. Again, Kant seems to be quite aware of this point when he declares the cause-effect relationship to be a category that falls under the heading of relation and is thus a primitive relational concept.[36]

If a causal power must thus be an indeterminate relation between a substance and the accidents that the substance brings about qua cause, one might still ask why a causal power could not become determinate through some further act. The crucial point here lies in Kant's understanding of the notions of activity and passivity involved in causality. Recall that for Kant a cause is not only constantly conjoined or even necessarily connected with its consequent effect, but also *brings about* or *produces* its effect by *actively determining* the boundary states of an object. The object so determined thus does not exist fully formed from the start with this determination, but rather passively receives it. Now, if one were to try to determine the cause (by some further causal connection), the result would, by parity of reasoning, be a determination of the substance that is the cause, not the activity essential to the causality of the cause. In other words, one could determine in this way a state of the cause, but not what is at issue, namely the very activity by which the cause brings about its effect.[37]

The crux of Kant's argument for the claim that the causality of the cause cannot become determinate is thus his idea that determinations are necessarily states of substances that are passive insofar as these states are brought about by the activities of causal powers of other substances.[38] In this respect, Kant's notion of a

---

[36] Another reason Kant has for not identifying power with substance is that it would force one to take what Kant thinks of as an overly restrictive view of faculties. One of Kant's most important immediate predecessors, Christian August Crusius, thought that human beings are endowed with understanding and will as faculties that are not reducible either to each other or to any more primitive power, a view that distinguished his position from that of the Wolffians. Though Kant seems to follow Crusius in thinking of the understanding and the will as distinct faculties, he is quite clear that reason (as active) and sensibility (as passive) are necessarily distinct (i.e., irreducible) faculties, for he famously criticizes both empiricists such as Locke and rationalists such as Leibniz for attempting to reduce the one to the other (at A271/B327). If faculties were simply epistemic powers, then the identification of power with substance would commit Kant to viewing human beings as composites of several substances. Though Kant may want to retain a kind of agnosticism about what we ultimately are (e.g., at the noumenal level), it is clear that he would not want to commit himself to the view that knowledge can occur only by means of the interaction within two distinct substances. As a result, Kant would have significant epistemological reasons for resisting the identification of powers with substances, at least in the case of human beings. As we shall see shortly, Kant thinks that in at least some cases substances should be thought of as containing only one basic power. There are several (mutually compatible) ways of resolving this tension. First, one could distinguish between different kinds of substances, e.g., those that are endowed with mental powers (i.e., faculties) and those that are not, and then hold that only certain kinds of substances can contain no more than one power. Second, one could maintain that even the claim that a substance can *have* only one power does not entail the *identification* of power with substance. Finally, one might think that the passages indicating Kant's commitment to the possibility of only one basic power may simply express his belief that, for methodological reasons, the identity conditions of the substances that underlie matter do not diverge from those of powers and that there would be no way to distinguish substances that would not also distinguish their powers.

[37] In principle, this point was already present in the Third Analogy. For what would make the circularity of mutual interaction vicious in the case where the causality exercised by each substance is understood to be a determinate state is the fact that each determinate state is supposed to produce the other. Without the determinacy and thus without the activity and correlative passivity, there would be nothing problematic about two states being related to each other.

[38] While it is standard to refer to substances as active and passive, there is a derivative sense in which one can speak of the active and passive "aspects" of substances in each case. The active "aspect" of a substance is the exercise of its causal power, while the passive "aspect" of a substance is the state (or determination) that the exercise of causal powers is responsible for.

determination is quite unlike contemporary notions of determinate states of affairs or properties, which would not normally be characterized as passive and which would thus not require a corresponding activity. What is Kant's justification for thinking of determinations as states that require the activity of a causal power? The most prominent justification for thinking of determinations in this way is found in Kant's Analogies of Experience. What drives the argument of the Analogies is the problem of time-determination, that is, the presupposition that the determinate temporal properties of objects are not immediately given along with the subjective order of our representations, but rather result only from a process of determination involving the relational categories. More specifically, the point is not that the objects already have determinate temporal properties and we simply must use the categories to determine the intuitions that are given to us in sensibility in order to *discover* these properties.[39] Rather it is that the objects do not already have determinate temporal properties.[40] The objects must be temporally determined by something else, something active (i.e., a determining ground), in order to have temporally determinate properties (29:818–19). It is thus appropriate that all three Analogies posit something active, namely some kind of determining ground, to bring about or determine the temporal features of objects, whether the activity takes the form of an inner sufficient ground for substance or an outer sufficient ground in the case of causality and mutual interaction.[41]

---

[39] Providing a detailed justification for a metaphysical reading of Transcendental Idealism extends beyond the scope of this paper.

[40] It may be helpful to clarify the senses in which the activity involved in the causality of the cause is and is not indeterminate. Such an activity is not indeterminate in the sense of being random or arbitrary. Nor is it for that reason atemporal, since, as we saw above, this activity is uniform and unchanging and thus unrestricted in its duration. Rather, it is temporally indeterminate in the sense that there is no intrinsic state of the substance that is responsible for the effect *and* to which one can assign a specific temporal index, since qua activity it cannot be a state that has been determined and hence made possible by an activity in the first place. (While it is true that Kant's position would still stand in stark contrast to Hume's even without understanding causal activity as temporally indeterminate in this sense, it does seem that he must in fact be committed to this further claim.)

One might attempt to distinguish between the *temporal* indeterminacy of the activity whereby a substance causes a change of state and the indeterminacy in the *kind* of cause it is and then object that granting the latter does not entail acceptance of the former, since uncertainty about what kind of cause a substance is does not entail that the substance does not exercise its causal powers (whatever they may be) at a determinate time. However, the inference in this case is not from our ignorance about the nature of a (fully determinate) cause to the indeterminacy of the temporal index of the cause, but rather from Kant's view that appearances have no determinate features independently of our knowledge of them and from his claim that temporal determinacy is derivative upon causal determination (i.e., the claim that one cannot attribute a specific temporal index to a substance except by attaching it to a state of the substance, which can occur only if there is a cause that brings that state about at a particular time).

[41] Further evidence that Kant understands causality in this way can be found in his discussion of the task of natural science. Kant asserts that "all natural philosophy occupies itself with the reduction of powers to a single basic power, which we cannot explain, namely that because something is, something else thereby follows. All basic powers must be given through experience" (29:772). Later in the same paragraph, Kant continues: "In natural science one has good reason to regard the attracting and repelling powers as primitive powers. Can there be in one substance many or only one basic power? For our reason there must be several because we cannot reduce everything to one, but the unity of each substance requires that there be only one basic power" (29:773–822, see also the *Metaphysical Foundations* 4:498–99). Though we shall return to this point below, Kant seems to think that in nature what we perceive are the effects of powers rather than the powers themselves. If one could perceive powers directly, one would not need to attempt to reduce powers to a single basic power. As a result,

Kant's reflections on causality in his Metaphysics Mrongovius lectures thus offer helpful explanations of important features of his model of causality in general and his notion of causal powers in particular. In addition to supporting the idea that Kant understands causality not in terms of events (or even determinations of substances) but rather in terms of causal powers, they reveal at a very abstract level that the relationship between cause and effect is irreducible and asymmetrical, that the asymmetry implies the determinacy of the effect and the indeterminacy of the cause, and that the asymmetry is derivative upon an active-passive dichotomy, since both the determinacy of the effect and the indeterminacy of the cause depends on the activity of a causal power.

## 4. ACTIVITY AND FORCES IN PHYSICS

Before we turn to the implications that Kant's model of causality might have for Kant's argument in the Second Analogy and for how Kant is replying to Hume, further consideration of some of its most basic features is in order. For even with the detailed description and justification of this model provided in the previous sections, one could, at this point, still press two further questions about it. First, very generally, given that this model has been described at such a high level of abstraction, can it be consistently applied to the concrete examples of causality that Kant explicitly developed? Second, more specifically, can we really make adequate sense of, or attach enough specific content to, the generic notion of activity by means of which an effect is brought about according to Kant's abstract model of causality? (This second question is made all the more pressing given that this activity is characterized as irreducible, asymmetrical, and indeterminate.) In order to address these questions, I consider how Kant fills out his abstract model of causality in two concrete instances: attractive and repulsive forces in physics and synthetic activities in consciousness. While the case of forces in physics provides a detailed illustration of Kant's abstract model of causality, it does not adequately clarify the notion of activity that is central to that model. By contrast, I argue, certain aspects of our conscious experience render intelligible Kant's notion of activity, even to those who are sympathetic to empiricist standards of intelligibility.

In his *Metaphysical Foundations of Natural Science* Kant devotes considerable attention to specific instances of causality in the realm of physics by discussing Newtonian attractive and repulsive forces at length. Specifically, in the Dynamics Kant argues that attractive and repulsive forces are necessary for a body to fill a determinate region of space, while in the Mechanics he argues that they are nec-

---

we must infer what kinds of powers there are from the kinds of observable effects they produce. But since we seem to have knowledge of powers in natural science only from their effects, and since powers are indeterminate given that, qua grounds, they are known only through their effects (i.e., their determinate consequences), we see, once again, that, as far as our knowledge can reach, a causal power is necessarily indeterminate. It is true that this particular argument is epistemological and thus consistent with the claim that the activity of causal powers might be indeterminate in an exclusively epistemological sense. However, Kant notoriously rejects the idea that the phenomenal world itself is (much less must be) fully determinate (in the Antinomies and in the Ideal of Pure Reason). Accordingly, it is certainly possible that the activity of causal powers is indeterminate in a metaphysical sense. Even without providing a full-fledged argument for understanding Transcendental Idealism in terms of a metaphysical notion of determination, one can still see room for such an interpretation.

essary for the communication of motion. In both the Dynamics and the Mechanics Kant holds that the respective masses of bodies are relevant to how strongly they attract and repulse each other. Further, in the Mechanics (4:542–43) Kant argues that the masses of bodies cannot change in the communication of motion.[42] Though Kant does not explicitly argue for the analogous point in the Dynamics (that is, that mass cannot change in the mutual attempts of bodies to penetrate and to resist the penetration of each other), a parallel argument can be constructed, since penetration and the resistance of penetration are simply special instances of the communication of motion (as Kant remarks at 4:536–37).

Kant's Newtonian theory of attractive and repulsive forces can thus be seen as filling out his abstract model of causality as follows. The first point to note is that Kant identifies bodies with spatial substances so that one can, speaking loosely, say that bodies are causes just as it is permissible to say that substances are. However, just as it is not, strictly speaking, substances per se that actually bring about an effect, but rather the causal powers that they have, so, too, it is not the bodies as such that are the cause of changes, but rather the attractive and repulsive forces that inhere in them. Moreover, the effect, which Kant understands generally as a change from one determinate state to another, is a change of some spatial state in the case of physics. For although attractive and repulsive forces have different kinds of effects and can be employed to explain different kinds of determinations, a fundamental similarity underlies these differences. It is true that attractive forces bring bodies closer together, while repulsive bodies move them farther apart, but this difference clearly pertains solely to the *direction* in which bodies are caused to move with respect to each other. Except for this point, their similarities dominate, since they both cause changes in the determinate spatial position of each other. Further, because bodies' masses do not change while determining the spatial properties of bodies, it is plausible to think that they form part of the essential nature of these bodies. Accordingly, what Kant says about attractive and repulsive forces fits in well with many of the general features of his abstract model of causality.

But it is striking that the parallels do not stop there. In Kant's account of attractive and repulsive forces one can find a plausible physical counterpart to the notion of activity that is central to his abstract model of causality. For at the level of physics, it is plausible to view the activity that produces a determinate change of state as the *exercise* of attractive and repulsive forces; whatever else an exercise of a force might be, it must be active. Moreover, the notion of the exercise of forces can even illustrate the three specific features that were attributed to Kant's notion of activity above, namely that it be irreducible, asymmetrical, and indeterminate. The exercise of these forces is irreducible because it cannot, so Kant believes, be reduced either to the motions that they produce or to the bodies in which they inhere. It is asymmetrical because the relata are very different; on the one side is a dynamical body (qua seat of forces) and its mass (i.e., an active spatial substance

[42] An important consideration that supports this point is Kant's claim in the First Analogy of Experience that "in all change of appearances substance persists, and its quantum is neither increased nor diminished in nature" (B224). While identifying the quantum of substance with mass may require argument, it is clear that Kant does accept such an identification.

along with its causal powers and essential nature), whereas on the other is, e.g., a change in motion (i.e., a determinate spatial event) that (passively) occurs. Finally, the exercise of a force is clearly not a determinate event or state of affairs that we could observe in the same way that we can a body's motion. One can see this last point even more clearly by considering that the exercise of a force is merely an "attempt" at bringing about a change of state. If two bodies attract and repel each other with equal force such that they are in a state of equilibrium and thus do not change their state with respect to each other, their forces continue to be exercised, that is, each body is still attempting to change the state of the other, despite the fact that these attempts do not result in a change of state due to the exercise of the counterbalancing force in the opposite direction. Because the exercise of attractive and repulsive forces is an irreducible, asymmetrical, and indeterminate activity, it displays the same structure as did the notion of activity involved in Kant's abstract model of causality.

In order to avoid misunderstanding how Kant's account of forces in physics might thus illustrate his abstract model of causality, it is important to recognize that the term 'force' is often used in a variety of ways, not all of which are consistent with Kant's view, if they are taken literally. For example, sometimes it is said that a body can transfer its force to another or that a body might exercise its force at one moment in time and not at another. From Kant's perspective, these ways of speaking require re-interpretation in line with his metaphysical account. If substances act continuously according to their unchanging natures, then, due to the analogy developed above between Kant's abstract model of causality and his account of forces in physics, it follows that bodies must exercise their forces continuously and uniformly in accordance with their mass. What changes is thus not the exercise of the forces themselves, but rather what *effects* the exercise of those forces will have. That is, a given substance acts in the same way at all times, but this activity can nonetheless cause different things to happen because the circumstances of the substance can be different (that is, because different substances can stand in different relations to each other). As those circumstances change, so will the effects that the substance brings about. Thus, two bodies located three feet apart will cause a greater reciprocal acceleration toward each other than those very same bodies separated by three light years. Their attractive forces are no different in the two cases; the bodies have exactly the same intrinsic and essential properties. What is different is the effect that those attractive forces have, and that difference is due to the different relations that the bodies stand in to each other. One might think of the constant and immutable "sphere of activity" that a substance has as being akin to a dynamical force field that spreads out over a certain region of space in a specific way and that causes objects to behave in different ways only to the extent different objects come to occupy different positions within the field. Accordingly, as long as one is careful to understand our various ways of speaking about forces in the appropriate way, one can identify a coherent interpretation of attractive and repulsive forces in physics in terms of Kant's abstract model of causality.

But even if we can provide a consistent interpretation of Kant's abstract model of causality in terms of the concrete example of attractive and repulsive forces in

physics, it still leaves open the second question that was raised above, namely whether we can really make proper sense of the generic notion of activity invoked in his abstract model. Though Kant reinterprets Newtonian attractive and repulsive forces in terms of his own metaphysics, he follows Newton (and other empiricists) in agreeing that the notion of activity that is expressed in the notion of an exercise is not something that we can directly observe (in intuition alone). Just as one does not literally see "the causality of the cause" or one billiard ball imparting motion to another, one does not see the exercise of attractive and repulsive forces. This point can be seen especially clearly in the case described above, where the forces of the two bodies counterbalance each other so that no change of state occurs, since in such a case there is no change to be seen and one must therefore infer the fact that forces are being exercised from other considerations. Since one can see nothing beyond the effects of these forces, there is no empirical content to physical forces *per se* (as opposed to their effects, which contain all such content).[43]

## 5. ACTIVITY AND SELF-CONSCIOUSNESS

At this point, it is helpful to consider how Kant's abstract model of causality can be illustrated not only by physical forces, but also by a specific kind of consciousness. For careful observation will reveal that what distinguishes activity in consciousness from the activity of forces in physics is the fact that we do have an immediate awareness of the self's synthetic activities, whereas we have no direct awareness of the exercise of Newtonian forces. As a result, our awareness of this specific synthetic activity will allow us to render intelligible Kant's generic notion of activity as it is employed in his abstract model of causality. Since providing an accurate description of the structure of self-consciousness is notoriously difficult, we must begin by carefully distinguishing different aspects of self-consciousness and pay special attention to the argumentative contexts in which Kant discusses self-consciousness, since they determine what aspect of self-consciousness is of interest to him in each case.

One must first distinguish between inner sense, on the one hand, and apperception, or self-consciousness, on the other. Though Kant does not always express his views on inner sense with perfect consistency, its primary systematic importance lies in its similarities and differences with outer sense.[44] Kant's general idea is that we can intuit external or spatial objects by means of outer sense and internal or non-spatial objects by means of inner sense. The result in either case is

[43] Kant is clearly aware of this point. As early as 1747 (in the opening paragraphs of the *True Thoughts on the Estimation of Living Forces*) Kant sees the dilemma one faces in describing forces. He explicitly criticizes those who describe forces, in this case moving forces, exclusively in terms of their empirically observable effects, namely the ability to cause motion, since such an explanation can appear vacuous. However, he is also aware that empiricists may not find forces intelligible. In fact, he objects to Leibniz's attempt to explain force in terms of "entelechia" for precisely that reason. Kant's solution in the *True Estimation* is to describe forces as being "active." Thus, Kant has already adopted the central feature of his Critical account in his first pre-Critical publication, though, at this point, he has no detailed explanation of what "activity" is, except in terms of its obvious contrast with passivity.

[44] For a discussion of some of the inconsistencies in Kant's doctrine of inner sense, see Henry Allison *Kant's Transcendental Idealism*, 255–71.

empirical knowledge (of either spatial or non-spatial objects), since both inner and outer sense provide intuitions, which are required to support claims to empirical knowledge. However, inner sense must also be distinguished from self-consciousness, because inner sense provides us with *knowledge* of the self, whereas apperception, which is not a source of intuitions, does not amount to knowledge (*Erkenntnis*) per se, but rather merely an *awareness* (*Kenntnis*) of the self.

In §24 of the second edition Transcendental Deduction Kant addresses the complex relationship between inner sense and self-consciousness indirectly by exploring how to resolve a paradox that arises from the very idea of knowledge of the self, namely that we must (paradoxically) be both active and passive with respect to ourselves in knowing ourselves. On Kant's general account of knowledge, if we are to have knowledge at all, then the object of knowledge must affect us to be *given* to us in intuition. That is, knowledge requires that we must be passive in some way, at least with respect to the object of our knowledge. This requirement poses no special difficulty in the case of knowledge of objects that exist externally to us since it is plausible that they can affect or act on us causally so that we can passively receive sensory information from them. However, in the case of knowledge of the self, matters are more complicated. Since the object of knowledge in such a case is not some external object, but rather the self, one is forced to admit that it is the self that acts on us so as to give us the relevant intuition. But this entails that in order to have knowledge of the self, the self must both act on and be acted on by itself, which, at least by Kant's lights, is paradoxical.

Kant argues that this paradox can be resolved by means of his distinction between inner sense and apperception. As we saw above, inner sense is the passive faculty through which objects affect us. The understanding, by contrast, is an active faculty in us by means of which inner sense can be acted upon or determined. Because inner sense and apperception are distinct from each other, one need not hold that self-knowledge requires one and the same thing to be both active and passive in the same respect. Accordingly, Kant's distinction between inner sense and understanding provides the basic framework for resolving the paradox of self-knowledge.

What is relevant about this resolution for current purposes is that in the course of explaining how the understanding can be an active faculty that affects inner sense, Kant explicitly claims that the self can be aware of its activity, e.g., in syntheses. Kant clearly states that the understanding "exercises that action on the **passive** subject, whose **faculty** it is, about which we rightly say that the inner sense is thereby affected," and its "synthesis, considered in itself alone, is nothing other than the unity of the action *of which it is conscious as such* even without sensibility" (B153, italics added). As we shall see in more detail shortly, earlier in the second edition Transcendental Deduction Kant also claims that a specific feature of *self-consciousness* requires that we be "conscious of synthesis" (B133). Since Hume consistently denies that he has any internal impression of a necessary connection or causal activity within himself, it is important to investigate this point with great care.[45]

[45] While Hume is quite explicit in this denial, it is significant that his second definition of causality in the *Treatise* requires that the idea of the cause "determines the mind to form the idea of the" effect (170). If "determines" is a causal notion and if one can be aware of it occurring, then Hume is implicitly committed to the thesis that we have an impression of causality in self-consciousness.

In §24, immediately after introducing the distinction between understanding (or apperception) and inner sense and claiming that we can be aware of the understanding's synthetic activities, Kant provides several illustrations in support of this claim:

> We cannot think of a line without **drawing** it in thought, we cannot think of a circle without **describing** it, we cannot represent the three dimensions of space at all without **placing** three lines perpendicular to each other at the same point, and we cannot even represent time without, in **drawing** a straight line . . . , attending merely to the action of the synthesis of the manifold through which we successively determine the inner sense, and thereby attending to the succession of this determination in inner sense. Motion, as action of the subject (not as determination of an object), consequently the synthesis of the manifold in space, if we abstract from this manifold in space and attend solely to the action in accordance with which we determine the form of **inner sense**, first produces the concept of succession at all. The understanding therefore does not **find** some sort of combination of the manifold already in inner sense, but **produces** it, by **affecting** inner sense. (B154–55)

All of the examples Kant describes here are cases in which one can be immediately aware of the understanding's activity in synthesis. While Kant explicitly casts doubt on whether one could represent a line at all without first drawing it, regardless of whether one follows him on this point, one can still see him maintaining a clear and tangible difference between i) drawing a line in thought and ii) a case where one first becomes aware of a point and then watches as that point moves across one's visual field in such a way that it seems to leave behind a series of points forming a continuous line. In short, in the first case, one is drawing a line, whereas in the second, one is merely watching as a line comes into existence in thought. It is plausible to interpret these cases as follows. In drawing a line in thought the understanding is *actively producing* the line and one can be immediately aware of this activity, whereas when one is watching a line being formed in thought, one is passive insofar as one is not directly aware of the creation of the line as depending on the understanding's activity (even should it be true that the line is created by such an activity).

The same kind of point is also present in Kant's distinction between motion as the determination of an object and motion as the action of a subject.[46] In the one case, there is an empirical object that changes its position in space and we represent that change as occurring in the object. In the other, there is the activity of the subject by means of which our representation of a change of position as such is produced, an activity that does not necessarily require (even the representation of) an externally existing object in motion. In a footnote meant to clarify this example, Kant interprets the latter case as involving the "**description** of a space" (B155), rather than a determination of an object. Whether one can determine that an object is in motion without also describing space, it is clear that there are two distinct points at issue. One concerns the attribution of motion to an object, while the other concerns what activity the subject is engaged in when it represents motion in space at all. It is in this latter case that we can be directly aware of our own activity in producing a representation of motion (just as we are when we draw a line in thought).

[46] Michael Friedman helpfully draws attention to this example in *Kant and the Exact Sciences* (Cambridge: Harvard University Press, 1992) 40, 131, and 200–1.

Further, Kant suggests that we are immediately aware of the understanding's activity in a wide range of cases and not just in what might appear to be the special cases described above. In a footnote to his discussion of the paradox of self-knowledge, he suggests that every "act of attention" (B156) is an example of us affecting inner sense. Again, the philosophical basis for Kant's point here can be made quite clear. It is one thing to have, or be conscious of, a certain representation. It is quite another to focus one's attention on it (or some aspect of it). Paying attention to a specific feature of our intuition is clearly distinct from simply having that intuition given through (inner or outer) sense, since one can have an intuition without paying attention to it at all. That is, it is plausible to think that the way in which one can, apparently at will, focus one's attention on one's own mental states should be understood as an act that we ourselves perform and can be aware of performing. Thus, whether one is drawing a line in thought or simply focussing one's attention on one's own mental states, it is clear that we are immediately aware of the understanding's activity, and this specific kind of self-awareness allows us to render intelligible the generic notion of activity that Kant employs in his abstract model of causality.

However, Kant's interest in our awareness of the understanding's activity extends beyond the value these particular examples have for his model of causality. For he attempts to build a more robust account of *transcendental self-consciousness* on the empirical awareness of activity illustrated by these examples. To see how Kant hopes to develop such an account of self-consciousness on this basis, one must first note that the various ways in which he discusses self-consciousness proper typically depend quite heavily on the argumentative context in which these discussions are situated. Kant's most famous discussion of self-consciousness occurs near the beginning of the second edition Transcendental Deduction where he addresses the issue of how we can explain the fact that various representations are mine. To this end, he describes apperception as "that self-consciousness which, because it produces the representation **I think**, which must be able to accompany all others . . . , cannot be accompanied by any further representation" (B132). A few lines later, he argues that the identity of apperception "contains a synthesis of the representations, and is possible only through the consciousness of this synthesis" (B133). The basic idea behind Kant's argument is that I can know that representations are mine only if I know that one and the same I has each one, but this can be known only if a) I connect them (given that they do not come into my consciousness already connected) and b) if I am aware that I am connecting them (since only my awareness of my connecting them allows me to know that each representation is being had by one and the same self).[47] Later in the Transcendental Deduction Kant will attempt to show that these connections must be represented by categories and that there is a tight argumentative link between such connections and knowledge of objects. In this way, Kant hopes to prove the objective reality of the categories, that is, that the categories are necessary for knowledge of objects.

[47] For a helpful reconstruction of this strand of Kant's argument, see Henry Allison *Kant's Transcendental Idealism*, 133–72, esp. 142–43.

While there is little consensus about much of Kant's argument in the Transcendental Deduction, there is widespread agreement about the idea that Kant is attempting to give an account of the fact that my representations are mine. And on this point, Kant appears to enjoy a significant advantage over Hume, who faces a serious dilemma.[48] Hume's empiricist principles seem to commit him to the claim that either we can know the self directly through an impression (just as we know any other object, whether external or internal) or we cannot know it at all. Hume avoids the former horn of the dilemma (by rightly noting that we do not have an impression of the self on a par with our impressions of other objects) only to succumb to the latter by asserting that the self is "a bundle" of perceptions without being in a position to explain what a bundle is (and why it is not a fiction in the way that bodies are). Hume himself famously admits (in the Appendix to the *Treatise*) that his bundle theory of the self cannot account for "the principles, that unite our successive perceptions in our thought or consciousness."[49]

Kant agrees with Hume that the first horn of the dilemma should be avoided, since he not only shares Hume's insight that we do not have an impression of a single, enduring self that remains the same throughout all the changes in our perceptions, but also develops this critical insight in detail in the context of his discussion of traditional metaphysical arguments. Kant expresses the point in his own words by noting (against rational psychologists such as Descartes and Leibniz) that we do not have an intuition of the self as a thinking substance (or subject). Yet Kant disagrees with Hume that one must therefore accept the second horn of the dilemma, for, as we have seen above, Kant thinks that we can be aware of the self and its identity *indirectly*, that is, by being aware of the activity of the self when it connects its various representations and by then inferring that it is one and the same self that does the connecting. By suggesting that the self can become aware of its identity not directly as an object of consciousness, but rather indirectly as the subject of activities of which we can be conscious, Kant is attempting to resolve the dilemma Hume faces.

At the same time, Kant does not beg the question against Hume by appealing to something that Hume could not in principle accept. It is true that Hume denies having any impression of necessary connection on the grounds that he has neither an external nor an internal impression of such a necessary connection. But if Hume's denial that we have an internal impression of necessary connection is based on the expectation that our internal impressions have to be exactly analogous to our external impressions (and then the fact that our internal impressions do not reveal anything similar enough), then Kant can be seen as pointing out that Hume's expectation is unfounded. Specifically, the analogy with external impressions may have led Hume to think that internal impressions would still be impressions of objects, and thus not to have been attentive to the features of consciousness to which Kant wanted to draw our attention. As a result, Hume could have noted the distinction Kant is making between a subject and an object of

[48] For an interesting discussion of Kant's reply to Hume on the issue of self-consciousness, see Patricia Kitcher "Kant on Self-Identity," *Philosophical Review* 91 (1982): 41–72.

[49] Hume, *Treatise*, 636.

consciousness and the way in which Kant explains how the activity of the self is necessary for one to become aware of the identity of the self in self-consciousness.[50]

Whatever merits Kant's account of self-consciousness might have with respect to the dilemma Hume faces, it is relevant to note that the structure of self-consciousness illustrates various features of Kant's abstract model of causality. The first point to note is that it is clear that the syntheses the self is aware of in self-consciousness are activities. Without these activities, there would be no connections between our representations. Specifically, these activities would seem to be instances of a particular kind of activity, namely an activity whereby a connection between representations is brought about as its effect. (Kant also states quite clearly that apperception *produces* "I think," which is obviously an activity of the self.) Moreover, the effect is determinate insofar as the particular kind of connection required by self-consciousness, namely a connection represented by the categories, *determines* our representations of an object (and its change of states) (cf. B128). In light of the fact, however, that determinacy is the result, product, or effect of these synthetic activities, the synthetic activities cannot themselves be determinate. We can thus discern a structure to self-consciousness that is similar in several respects to what we saw in Kant's abstract model of causality.[51]

There is yet another aspect of self-consciousness, distinct from the synthetic activity required to explain how I can know that my representations are mine, that is also relevant, though in a different way, to understanding Kant's model of causality. In the Paralogisms, where Kant rejects various arguments that rational psychologists attempt to offer on the basis of self-consciousness alone, he explores the peculiar structure of self-consciousness in detail (e.g., at B423 fn.), arguing for an asymmetry similar in certain respects to the asymmetry that a cause bears to its effect.[52] In this context, Kant not only distinguishes between the object and subject of consciousness, but also notes that our awareness of the subject in self-consciousness is radically different from our awareness of objects.[53] It is one thing

---

[50] Had Hume understood self-consciousness in this way, he might have been led to revise other parts of his position as well (e.g., his epistemology, his general account of causation, etc.).

[51] One might object (e.g., on behalf of Leibniz) that because Kant must rely on consciousness to render intelligible the notion of activity, he cannot obviously or immediately claim that this same notion of activity is also at work in physical cases, especially if Kant admits to important differences between the two cases. Perhaps, the objection continues, physical cases must be unintelligible, or perhaps they stand in need of a deeper metaphysical explanation (e.g., in idealistic terms). In response, I would claim that the activity we are aware of in our own consciousness is supposed to be simply a *specific* instance of a *generic* notion of activity that Kant wants to employ in his general model of causality. That is, the objection provides no reason to think that Kant *cannot abstract* from those particular features of consciousness that happen to attend the activity we can be directly aware of in self-consciousness.

[52] For helpful discussions of this aspect of Kant's account of self-consciousness, see Karl Ameriks "From Kant to Frank: The Ineliminable Subject," in *The Modern Subject: Conceptions of the Self in Classical German Philosophy*, ed. K. Ameriks and D. Sturma (Albany: SUNY Press, 1995), 217–30, and Manfred Frank "Is Subjectivity a Non-Thing, an Absurdity [Unding]? On Some Difficulties in Naturalistic Reductions of Self-Consciousness," in *The Modern Subject*, 177–97.

[53] Berkeley encounters a similar difficulty on this point, since for him ideas are necessarily passive, yet the self is supposed to be active. He solves the problem by claiming that he has not an idea, but rather a "notion" of the self. While introducing a notion of the self in this fashion might initially seem to be *ad hoc*, Berkeley may be responding to an independent philosophical concern. For an excellent discussion of Berkeley's position, see Robert Adams "Berkeley's 'Notion' of Spiritual Substance," *Archiv für Geschichte der Philosophie* 55 (1973): 47–69.

to say that we do not see the self in the same way as we do spatio-temporal objects. It is another thing altogether to say that we cannot even *describe* it in any positive way by using the predicates that we use in describing objects as it makes the use of such predicates possible in the first place.[54] On Hume's view, there is presumably nothing special about the self that precludes our seeing it. Rather, it simply does not exist (as a single, identical entity) and its non-existence accounts for the fact that we do not see it. On Kant's account, there is something very special about the self or at least about our mode of conscious access to it. Since it is not a directly observable object, but rather, insofar as we can be aware of it at all, issues in an activity rather than a determinate state of an object, it is not something that can be described by using the same concepts that apply to the determinate states of external objects that we observe in exactly the same way. According to Kant, therefore, Hume mistakes our inability to describe the self in typical object-language terms with an inability to be aware of it at all.

At this point, we have seen how Kant's Newtonian account of attractive and repulsive forces in physics and his novel explanation of empirical and transcendental self-consciousness illustrate several basic features of his abstract model of causality. The case of physical forces was especially well-suited to providing an intuitive picture of the basic structure of Kant's general model, while the case of self-consciousness proved to be of indispensable aid in providing clear and intelligible content to the notion of activity that Hume (or any other standard proponent of empiricism) might have found most controversial about Kant's model of causality. In short, we now have in front of us both general and specific descriptions and thus a comprehensive account of the basic features of Kant's model of causality.

## 6. THE UNIFORMITY OF NATURE

Let us now return briefly to the argument of the Second Analogy to consider how it might naturally be understood in terms of the model of causality described above. The basic thrust of the Second Analogy is to argue that causality is necessary for the determination of successive states in an object, since objects do not have temporal determinacy independently of such causal determination (e.g., merely as given to us in intuition). But if the model of causality described above dictates that causal determination is to be understood in terms of grounds, then the point of the Second Analogy is that change must be explained in terms of grounds. Yet since Kant argues that the ground that causes change cannot itself be changing at the same time—at least not if one is to avoid an infinite regress—change presupposes an unchanging ground. The idea of unchanging grounds might be thought sufficient, however, to support causal laws, since the grounds cannot change in the future and thus must, it would seem, bring about the same effects as before.

---

[54] For perhaps different reasons, Shoemaker endorses both of these points in "Self-reference and Self-awareness," *The Journal of Philosophy* 65 (1968): 555–67, esp. 563, when he says: "I think that the main source of trouble here is the tendency to think of awareness as a kind of perception, i.e., to think of it on the model of sense-perception" and when he notes that if the use of first-person pronouns as a subject "were not possible then there would be much else, and much that we take for granted, that would also not be possible" (564).

One might object to his final inference as follows. The argument of the Second Analogy, so understood, establishes only that a ground cannot change *while it is determining successive states in an object*, and not that a ground cannot change *at all*,. But if a ground can change, then it would seem possible that causal laws cannot be established, since they are warranted only if grounds are immutable.

Yet Kant might have had at least two distinct reasons for thinking that this objection could be overcome. First, Kant might take recourse to an idea that that he had developed in his Inaugural Dissertation, namely that substances' natures are general, in order to establish that the grounds constituting them cannot change. The idea would be that if a substance's nature is truly general, then it will hold not only for any substances that are part of the world to which that substance happens to belong, but also for all times, that is, for all states of all such substances. Accordingly, if a ground were to change at some point between $t_1$ and $t_2$, then the generality of the nature would be compromised. Accordingly the generality of natures (or of the grounds that form them) might entail unchanging causal laws.

Second, if grounds are to be understood as involving indeterminate activities rather than determinate states, then Kant might also have thought that grounds are not capable of change (at least not in the way that determinate states are). For Kant understands change as a change of determinations, and if grounding activities are not determinations, then whatever might happen to them could not be represented as a change of determinations. The obvious objection to this line of thought is that it seems arbitrary to restrict change to change of determinations. Why not allow change of determinings in addition to change of determinations? In light of the indeterminacy of the activity involved in grounding, it is unclear how such change is to be understood other than in terms of changes of the determinate states that the grounds produce. And if a change of grounds can be understood only in terms of changes of the determinations they produce, then we would be faced with the issue of what the identity conditions of grounds (or rather groundings) are. Given Kant's conception of grounds, one can see why it would not be tempting to view them as capable of change. For if grounds cannot be directly perceived, then the primary basis for asserting that they are changing is removed as well, at least as long as a different interpretation is available. For if one accepted changing grounds, then one would have to undertake an impossible task, namely explaining why grounds changed in precisely this way at precisely this time, and whatever explanation one gave, it would, so it seems, have to be in terms of further grounds that either changed or did not, in which case no real explanatory progress would have been made. Thus, instead of saying that a ground has changed, it seems much more attractive to assert that a different ground is active in bringing about different effects.

Does this reconstruction of Kant's argument not prove too much? That is, would the argument, if successful, not establish that the sun will necessarily rise tomorrow, and would this not be objectionable given that we could imagine events that would prevent this from happening? In numerous passages Kant makes it clear that empirical causal laws can be determined only on the basis of experience (and that regulative principles may be indispensable in discovering such laws). But if I see that one determination follows another in a certain set of cir-

cumstances, does that not immediately establish a causal law stating that the one determination follows another in such circumstances, such as the sun rising in the morning? The point to Kant's claim and to the notion of grounds that he believes supports that claim is that his complex notion of grounds and the model of causality based on it merely supply a *formal ontological framework* that must be filled in with empirical content.[55] Accordingly, this framework entails only that whatever grounds and causal laws have held in the past will not change in the future. Thus, even if Kant were to establish the necessity of grounds (and the causal laws that depend on them) for the determination of the changes that occur in the world, the epistemological question of ascertaining what grounds and laws exist in the world has not been addressed at all. As Kant clearly indicates, this epistemological question can be answered only by consulting experience and can presumably never be established with absolute certainty, since one can never rule out the possibility that future evidence might require a revision in our understanding of what grounds exist in nature. In short, knowing that there are immutable grounds does not at all resolve the question of what empirical content they have.[56]

As a result, Kant's model of causality helps us to make sense of why Kant might have thought that his argument in the Second Analogy could prove the existence not only of necessary connections but also of causal laws. For Kant's model of causality is based on unchanging grounds, and unchanging grounds entail, so Kant thinks, that substances will continue to act in the future as they have in the past. Moreover, not only does our focus on Kant's model of causality expose what Kant's reasoning for the necessity of causal laws is, but it also explains why he would have moved back and forth between stronger and weaker causal principles without any apparent concern for the differences between them. Since one and the same model of causality provides the resources that can establish both principles, it is natural for him to switch back and forth.

## 7. KANT'S REPLY TO HUME

Given this detailed understanding of Kant's model of causality, we can now return to the question with which we began: How Kant is replying to Hume on the issue

[55] For a discussion of how a formal (mathematical) framework could be filled in with mechanical and physical content, see Michael Friedman, *The Dynamics of Reason* (Stanford: CSLI Publications, 2001).

[56] Even if this interpretation does accurately represent Kant's intentions, it is unclear that Kant's arguments can successfully carry the weight of the strong reading of the Second Analogy at a metaphysical level. For even if the natures and grounds of substances *might* be general in the full-blooded sense that would be required to support causal laws, Kant has given no explicit argument for thinking that they *must* be general in such a rich sense. Also, however plausible it may (or may not) seem to assert that permanent grounds cannot change, Kant has given no argument for the identity conditions of grounds that would definitively determine that permanent but changing grounds are impossible. Finally, even if one were to grant both that indeterminate activities cannot change in determinate ways and that we have no cognitive means by which indeterminate changes might be known, that still does not provide a full justification for claiming that indeterminate activities could not change nonetheless. In short, although one can now understand why Kant may have been tempted to think that his model of causality would entail that not only causality, but also causal laws would be required in order to have knowledge of objective succession, it is also clear that he does not develop fully explicit arguments that would in fact suffice to establish the Second Analogy's stronger principle.

of causality? Before we investigated their respective models of causation, Kant's reply to Hume seemed, at least in principle, perfectly straightforward. For it appeared that Kant was simply affirming what Hume was denying, namely that a cause is necessarily connected with its effect, and the question was simply whether Kant's argument, properly understood, could refute Hume's position. If the Second Analogy can actually establish that necessary connections are required for our knowledge that one state follows another and if the considerations raised in the previous section about the nature of grounds that were involved in explaining Kant's model of causality show why Kant might have thought that such an argument could be used to establish the necessity of causal laws as well, then one might easily think that investigating Kant's model of causality has had the perhaps unexpected benefit of showing that Kant's reply to Hume is both straightforward—in line with our initial expectations—and comprehensive, since Kant's arguments can establish *both* of the principles that Hume's skeptical arguments called into question.

However, this interpretation of Kant's reply to Hume fails to appreciate that the understanding of Kant's model of causality afforded by our detailed discussions above reveals that Kant's argument in the Analogies does not so much *refute* Hume's skeptical arguments as it does *beg the question* against Hume by assuming a model of causality that Hume would reject. For what Kant takes an effect to be, namely a continuous change from one determinate state to another, is not what Hume means by an effect, namely a determinate state of an object at a particular moment in time. Nor is the cause for Kant—an indeterminate activity of a substance in accordance with its ground that determines the change of state in the effect—identical to the cause for Hume—a determinate state at a particular moment in time (that is regularly followed by another determinate state). Thus, when Kant asserts that causality is necessary for knowledge of an effect, he assumes that there is a continuous change from one determinate state to another that is caused by a continuous, but indeterminate activity of a substance, whereas Hume is committed only to one state at one moment in time being followed by another such state (and to states similar to the first being following by states similar to the second on other occasions). As a result, Hume's denial and Kant's assertion of a necessary connection between cause and effect do not, in fact, contradict each other, because they are talking about the possibility or impossibility of a necessary connection between completely different kinds of entities. Accordingly, the model of causality that Kant is working with is simply not the same as Hume's and it is open to Hume to reject the very starting point of Kant's argument on the grounds that he does not share his most fundamental assumptions.

One natural strategy at this point would be to attempt to reinterpret Kant's model in such a way that Kant's argument could be made intelligible on Humean terms. On this strategy, while Hume expresses skeptical doubts about a necessary connection *between* cause and effect, as he understands these terms, Kant could reinterpret Hume's view as raising doubts about such a connection *within* the effect, as Kant understands the term, since it is only within the effect that there is the kind of temporal asymmetry that was required in Hume's account. If Kant could show that the first state of the effect is necessarily connected with the sec-

ond state of the effect, then he might be able to reply to Hume directly and on Hume's own terms. For in that case he will have established a necessary connection where Hume thought that none could be found.

However, recall how vast the differences are between Hume's and Kant's models of causality and between the more general ontologies of which their models are a part. Hume's events are states of affairs at instantaneous moments in time, whereas Kant's events are continuous changes of state over time. Hume accepts only events and contingent relations between them, while Kant accepts substances (both noumenal and phenomenal) as well as a series of indeterminate ontological posits such as inherence, grounds, causal powers and "the causality of the cause" that cannot be understood at all in terms of Humean events. Hume attempts to construct an account of the world solely on the basis of such discrete events, whereas Kant is concerned to explain that the relational categories are necessary for us to know the temporally determinate states of an object that belongs to a single spatio-temporal world. Hume's events are distinct from each other, while Kant attempts to establish (the necessity of) grounding or dependency relationships between substances and their determinate states. Hume's events are neither active nor passive, Kant's causes are active and his determinations passive. It should thus be clear that in spite of the fact that Hume and Kant are both attempting to provide an account of one and the same objective world, the causal models and the basic ontologies that they invoke to describe the world have nearly nothing in common.

But if Hume's and Kant's models of causality and basic ontologies have almost nothing in common, it follows that Kant's argument cannot be understood as a refutation of Hume's position. For the kind of reinterpretation suggested above presupposes that they share at least some basic terms, terms that could allow one to eliminate their apparent differences. Kant's assertion that causality is necessary for (knowledge of) objective succession or that mutual interaction is necessary for (knowledge of) objective simultaneity presupposes temporally indeterminate grounds that determine the relevant temporal states of objects, a presupposition that Hume, in effect, rejects by accepting only fully determinate events as his fundamental building blocks, since they must exist at a particular instant in time. Therefore, Kant's Analogies of Experience immediately beg the question against Hume and thus cannot be used to refute him.

If Kant cannot refute Hume, must one deny that Kant has any reply to Hume at all (and that his remarks about Hume in the *Prolegomena* and elsewhere are simply a somewhat misleading rhetorical device)? First, nowhere does Kant ever use the term "refute" in discussing his reaction to Hume, so we are not forced to discount any direct textual evidence by acknowledging that Kant is not refuting Hume.[57] Moreover, if one takes Kant in his proper historical context, one can see more clearly that Hume is important to Kant not because Hume's position stands in need of refutation, but rather because Hume develops provocative critical in-

[57] At A195–96/B240–41 Kant does suggest that his view "contradicts" a consistent empiricist analysis of causality, but I take it that in this passage Kant is simply expressing the radical *difference* between what his view can account for, namely necessity, and what the empiricist can have, namely something that is "merely feigned" and lacking "true universal validity."

sights into a variety of particularly fundamental arguments and issues, insights that Kant wanted to accommodate within his own system. Thus, Kant saw quite clearly the power of Hume's criticism of Cartesian arguments for thinking substances, his attack on the notion of substance in general, and his broad assault on traditional theistic arguments. But notice that in each of these cases Kant took Hume's criticisms not as reasons for abandoning his own position nor as criticisms that needed to be refuted, but rather for rethinking what *kind* of arguments would establish the positions he would continue to hold. Instead of accepting purely dogmatic, metaphysical arguments that must appear highly contentious, at best, and quite feeble when faced with Hume's insightful criticisms, Kant saw the need to develop epistemological arguments that would make explicit what cognitive powers and activities are required for knowledge of the world as he understands it to be possible.

Viewed in this light, one can see that Kant's reaction to Hume on the issue of causality is no different. No doubt, Hume pointed out quite forcefully that one cannot *immediately perceive* necessary connections in the world or any causal laws that would be based on them. But Kant took the point of these powerful considerations to be not that one must give up on necessary connections or causal laws, but rather that one must argue for them in a novel way. Specifically, Kant sees that he can exploit the dependence of our knowledge of temporal determinacy (in the form of succession and coexistence) on specific kinds of causal structures (specifically, causality and mutual interaction) to show how our experience (or knowledge) of a single spatio-temporal world is possible. The result is that Kant develops both a specific model of causality and a general approach to philosophy that do not refute Hume's position, but rather compete against it.[58]

While there is no doubt that discovering and developing a fundamentally new philosophical methodology represents a significant intellectual accomplishment, what makes Kant's achievement all the more impressive is the way in which he displays sophistication and subtlety (even if not immediate clarity or popularity) in applying that method to address long-standing philosophical problems, in this case, the problem of causality. Note how his model of causality avoids several one-sided positions. Instead of trying to reduce causality to a purely contingent relation between events or *simply* invoking substances and causal powers, Kant argues for an intermediate position by insisting on the idea that substances exercise their causal powers by means of an asymmetrical activity of determination in bringing about or producing events according to their natures.

Moreover, by emphasizing the activity of determination in this way, Kant is in a position to incorporate this notion into his distinctive philosophical system in a variety of ways. For one, he can use the notion of an indeterminate activity to account for the way in which Newtonian attractive and repulsive forces are sup-

[58] If Kant is competing against, rather than refuting Hume, the question immediately arises as to how one should attempt to decide between them. But note that because of the magnitude of the differences between Kant's and Hume's views on causality and the overall projects of which they are a part, it is far from obvious that one can identify neutral "internal" criteria that would be able to determine who would win such a competition, and it is not obvious that we have any access to an external standpoint or a "view from nowhere."

posed to be *exercised* in the interactions of bodies. For another, it can be used describe certain aspects of (the unity of) self-consciousness in his philosophy of mind. Finally, because Kant makes the notion of the activity of determination central to his metaphysics, one can understand, at a very abstract level, that he might want to understand autonomy—which may be attractive for independent reasons stemming from the metaphysics of morality—in terms of self-determination, since the core metaphysical notion of determination would then stand in need of no additional explanation (even if complex problems concerning the self and the peculiar kind of legislation it is involved in do require extensive elaboration). This is neither to endorse any of Kant's arguments nor to suggest that his philosophical system is preferable to Hume's (or anyone else's). It does, however, give at least some indication of the power and virtues of Kant's position when the full range of resources to which it can appeal are appreciated.[59]

[59] I am indebted to Robert Adams, Karl Ameriks, Nancy Cartwright, Daniel Garber, Rick Grush, Patricia Kitcher, Rudi Makkreel, Wayne Martin, Sam Rickless, Don Rutherford, Tad Schmaltz, audiences at the University of California at San Diego, Yale University, Duke University, University of Iowa, Emory University, and the Midwest Seminar in Early Modern Philosophy (University of Wisconsin-Milwaukee, December 1999), and an anonymous referee for the *Journal* for helpful comments on or discussion of earlier versions of this paper.